



Universidad Nacional del Litoral  
**Facultad de Ingeniería y Ciencias Hídricas**

## **ESTADÍSTICA**

**Ingenierías: Recursos Hídricos-Ambiental-  
Agrimensura**

---

### **TEORÍA**

*Mg. Ing. Susana Vanlesberg*

**Profesor Titular**

**UNIDAD 1**

**Introducción-Probabilidad**

## INTRODUCCIÓN

Hablar en la actualidad del conocimiento estadístico es decir que es necesario no sólo para ser un buen y eficiente ciudadano sino también para una toma de decisiones efectiva en el área en que le toque actuar. La creciente importancia de la *Estadística* en distintos campos es un hecho real y concreto tal el caso por ejemplo de su aplicación en producción, biología, economía, política, meteorología, cartografía y por supuesto hidrología e hidráulica.

Casi todos los días utilizamos los conceptos estadísticos en las diversas facetas de nuestra vida.

Por ejemplo un candidato a un cargo en el gobierno o en el senado quisiera saber el porcentaje de electores que lo apoyarían. Tiene varias formas de poder hacerlo: podría encargar a su equipo que llame por TE a todas las personas que integran el padrón de la provincia a la que pertenece y que pregunten por quien votarían; puede salir a las calles detener a 10 personas en condiciones de votar y preguntarle su opinión; puede seleccionar una muestra de 2000 electores de la provincia hablar con ellos y basándose en esta selección, realizar un cálculo del porcentaje de personas que lo votarían. En el curso aprenderemos a poder decidir que la última opción es la mejor.

¿Cómo se define la palabra *estadística*? Es un término que encontramos con frecuencia. En el lenguaje más común se refiere a información numérica, como ejemplos número promedio de automóviles marca Ford vendidos en el último mes en la ciudad, el % de alumnos que culminarán su carrera este año en la Facultad, el número promedio de días de lluvia en el mes de julio, la extensión promedio de los picos de contaminación en un tributario al Paraná medidos en un punto de control desde el año 2000 es de 8 meses.

Todos los ejemplos anteriores son *estadísticas*. Un conjunto de información numérica se conoce como *estadísticas* (plural). Generalmente se utiliza esta información para hacer gráficas y llamar la atención del lector.

**La ciencia Estadística va mucho más allá de la mera recopilación y publicación de información numérica, se define como la ciencia que se ocupa de recolectar, organizar, presentar, analizar e interpretar datos para ayudar a una toma de decisiones más efectiva.**

*¿Por qué estudiar estadística?*

Si se revisaran los programas de distintas carreras se vería que la educación estadística aparece en casi todas ellas. ¿Por qué? ¿Cuáles son las diferencias entre un curso de estadística en una carrera de Psicología, en una Ingeniería, en una carrera de administración etc? En general los temas son bastante similares, las diferencias estarán en las aplicaciones o casos en los que se apliquen las técnicas estadísticas y en el nivel de matemática exigida según la profundidad con que se los enseñe.

**La primera razón para estudiar estadística en todas las carreras es que nos encontramos con información numérica.**

**La segunda razón es que las técnicas estadísticas se utilizan para tomar decisiones que afectan la vida diaria y la profesional.**

**La tercera razón es que el conocimiento de los métodos estadísticos ayuda a entender por qué se toman ciertas decisiones y aporta una mejor comprensión respecto a la forma en la que nos afectan esas decisiones.**

Sin importar el tipo de trabajo que se tenga siempre habrá que enfrentarse con la toma de decisiones para lo cual una comprensión del análisis de datos será de gran ayuda.

Los métodos que forman parte de la *Estadística* permiten trabajar datos, los cuales han sido obtenidos de forma repetitiva. Por lo tanto no debe extrañar el hecho de su amplia aplicación profesional. La mayor parte de los trabajos que se realizan en las áreas de recursos naturales se relacionan con el análisis de datos y toma de decisiones por lo tanto la *Estadística* constituye una importante herramienta con la cual cuenta el profesional.

Los datos geofísicos pueden ser muy variados pero todos presentan un *rasgo* común que se deriva de circunstancias que están afectadas por la *casualidad*; es decir en todas las situaciones existe la presencia de efectos que no se pueden predecir porque son el resultado de factores que no se pueden controlar y muchas veces ni siquiera ser enumerados; esto es lo que comúnmente se denomina *aleatoriedad*.

Un profesional que planifica y diseña generalmente analizará eventos futuros cuya magnitud y tiempo de ocurrencia no son conocidos con certeza; en general fenómenos tales como precipitaciones, caudales, evaporación, imágenes satelitales, etc. son el resultado de la combinación de eventos más pequeños (por ejemplo en la formación de la precipitación intervienen una serie de procesos que involucran a las leyes de la termodinámica e hidrodinámica). Muchos procesos pueden ser analizados en forma determinística pero en general se hace necesario recurrir a descripciones en términos estadísticos.

**El uso de los métodos estadísticos posibilita la síntesis de la información disponible a algunos valores que sugieren claramente la naturaleza de los datos crudos y además brindan un procedimiento para la toma de decisiones.**

Normalmente se puede disponer de un número limitado de observaciones de la variable de interés, tal número de observaciones es lo que se denomina *muestra*; ésta es extraída de un conjunto mayor denominado *población*, que es el conjunto de todas las posibles observaciones que se pueden obtener relacionadas con el fenómeno en cuestión.

En muchos casos la población es demasiado grande como para trabajar con ella ya que se necesitaría mucho tiempo y dinero para su relevamiento completo, por lo tanto el rol de la muestra es importantísimo para obtener a partir de ella conclusiones válidas acerca de la población de la cual proviene. Por lo tanto el material básico para cualquier análisis estadístico lo constituye la muestra.

Es interesante observar que, al menos desde la época del Imperio Romano hace 2000 años, ya existían recopilaciones de datos. Pero la introducción de los métodos estadísticos en las ciencias físicas, sociales, médicas etc., comenzó hace sólo unas cuantas décadas.

## CONCEPTOS BÁSICOS

Es frecuente en ingeniería realizar simplificaciones tales como no considerar la fricción, adoptar un fluido ideal, y esto con el fin de obtener modelos matemáticos relativamente sencillos.

Generalmente estos modelos son determinísticos, es decir un número describe a cada variable independiente y una fórmula predice un valor específico para la variable dependiente. Para el mismo conjunto de valores de la variable independiente se obtendrá siempre el mismo conjunto de variables dependientes.

Cuando lo anterior no sucede es decir no se obtienen los mismos resultados para el mismo conjunto de valores de la variable independiente es porque hay algo que no se puede explicar y eso es lo que se denomina incertidumbre, esto puede ser por variación inherente o por conocimiento incompleto del fenómeno, los modelos que se derivan son probabilísticos y se los analiza haciendo uso de la **Teoría de Probabilidad**.

Por ejemplo no se puede afirmar que un día cualquiera lloverá, y mucho menos decir que la precipitación alcanzará cierta cantidad de milímetros. En cambio, es posible decir que es probable que mañana llueva, con lo que se introduce el concepto de probabilidad.

**La teoría de probabilidad está formalmente relacionada con los experimentos y sus resultados, donde el término "experimento" está usado en un sentido general.**

### Características de un Experimento Aleatorio:

- El experimento se puede repetir indefinidamente bajo idénticas condiciones.
- Cualquier modificación a las condiciones iniciales de la repetición puede modificar el resultado.
- Se puede determinar el conjunto de posibles resultados pero no predecir un resultado particular.
- Si el experimento se repite gran número de veces entonces aparece algún modelo de regularidad estadística en los resultados obtenidos.

El conjunto de todos los resultados posibles de un experimento que pueden identificarse antes de realizarlo se denomina **espacio muestral**, generalmente indicado con la letra **E**. Cada uno de sus elementos se denomina **punto muestral**.

De acuerdo a los valores que el espacio muestral contenga, puede ser clasificado como:

**-Finito:** si contiene una cantidad dada de valores posibles. Por ejemplo medir la dirección del viento en un aeropuerto: valores posibles identificados a priori entre  $0^\circ$  y  $360^\circ$ .

**-Infinito Numerable:** si se sabe que contendrá una gran cantidad de elementos pero posibles de enumerar, y pueden ponerse en correspondencia con el conjunto de los enteros positivos. Por ejemplo: número de días de lluvia en un período de 10 años en la ciudad de Santa Fe.

**-Infinito o continuo:** sus elementos se corresponden con los números reales. Por ejemplo: considerar el valor de los errores en mediciones de distancias, contenido de un contaminante en muestras de agua.

**Evento o Suceso** es un conjunto de puntos muestrales del espacio muestral  $E$  correspondiente a un experimento aleatorio.

Diremos que **ocurre o se presenta el suceso** cuando al realizarse el experimento aleatorio, da lugar a uno de los resultados elementales pertenecientes al subconjunto  $A$  que define el suceso.

Se pueden considerar cuatro tipos de sucesos según el  $N^\circ$  de elementos que lo formen:

- **Suceso elemental, suceso simple o punto muestral** es cada uno de los resultados posibles del experimento aleatorio, entonces los sucesos elementales son subconjuntos de  $E$  con sólo un elemento.
- **Suceso compuesto** es aquel que consta de dos o más sucesos elementales.
- **Suceso seguro, cierto o universal** es aquel que consta de todos los sucesos elementales del espacio muestral  $E$ , es decir, coincide con  $E$ . Se le denomina seguro o cierto porque ocurre siempre.
- **Suceso imposible** es aquel que no tiene ningún elemento del espacio muestral  $E$  y por tanto no ocurrirá nunca. Se denota por  $\emptyset$ .

El **suceso complementario** de un suceso es aquel formado por los elementos del espacio muestral que no están incluidos en el suceso.

Al observar el fenómeno *velocidad de viento* en cierta localidad, el espacio muestral correspondiente se puede graficar como se ve en la Fig. N° 1, correspondiendo  $0^\circ$  al Norte,  $90^\circ$  al Este,  $180^\circ$  al Sur,  $270^\circ$  al Oeste. La zona sombreada más oscura representa al evento en el cual las velocidades están entre 100 y 200 km/h, y el sentido es Norte-Sur. El evento complementario está dado por las velocidades y sentidos que se encuentran en la zona sombreada más clara.

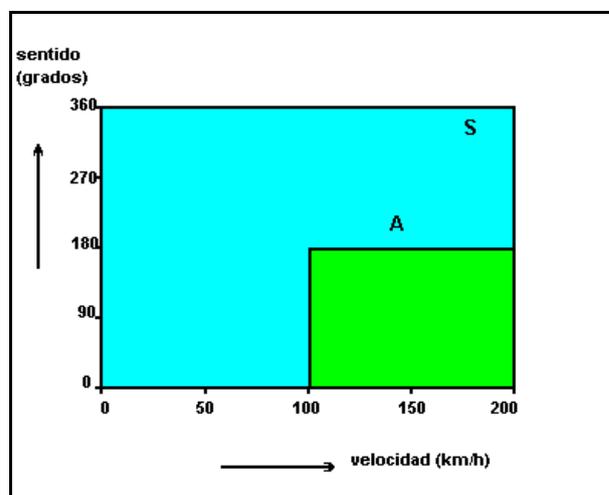


Fig. N° 1

## OPERACIONES ENTRE EVENTOS

Los eventos pueden estar relacionados de varias formas:

-El conjunto de puntos muestrales comunes a dos eventos se denomina *intersección*.

En general, dados  $n$  sucesos  $A_1, A_2, A_3, \dots, A_n$ , su intersección es otro suceso formado por los resultados o sucesos elementales que pertenecen a todos los sucesos  $A_i$ .  $\bigcap_{i=1}^n A_i$ .

-Si dos eventos A y B no contienen puntos elementales en común, se dice que son *mutuamente excluyentes* ( $A \cap B = \emptyset$ ); esta relación puede extenderse a más de dos eventos  $\bigcap_{i=1}^n A_i = \emptyset$ .

-La unión de dos eventos A y B es el evento que contiene los puntos muestrales pertenecientes a uno o a otro o a ambos ( $A \cup B$ ).

En general, dados  $n$  sucesos  $A_1, A_2, A_3, \dots, A_n$ , su unión es otro suceso formado por los resultados o sucesos elementales que pertenecen al menos a uno de los sucesos  $A_i$ .  $\bigcup_{i=1}^n A_i$ .

En la Fig. N° 2 se pueden observar la unión y la intersección. El conjunto de velocidades comprendidas entre 100 y 120 km/h y sentido Este-Sur es la intersección y el evento formado por las velocidades comprendidas entre 100 y 200 km/h y sentido Norte-Sur o velocidades entre 50 y 120 km/h y sentido Este-Sur es la unión.

### Sistema Exhaustivo de Sucesos

Dados  $n$  sucesos  $A_1, A_2, A_3, \dots, A_n$  de un experimento aleatorio:

Éstos forman una **colección o sistema exhaustivo de sucesos** si la unión de todos ellos es igual al espacio muestral  $E$ .

$$A_1 \cup A_2 \cup \dots \cup A_n = \bigcup_{i=1}^n A_i = E$$

Forman un **sistema completo de sucesos o una partición de  $E$**  si, además de la anterior condición, se cumple que son disjuntos dos a dos, es decir, son mutuamente excluyentes, disjuntos o incompatibles.  $A_i \cap A_j = \emptyset \quad \forall i \neq j$

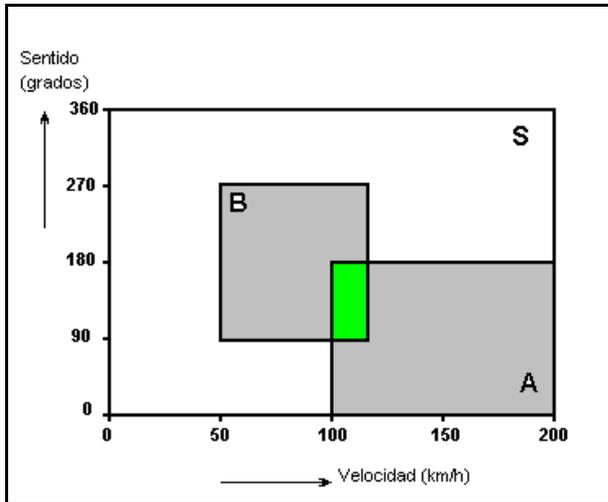


Fig. N° 2

El conjunto de todos los sucesos elementales que constituyen un espacio muestral forman una colección de **sucesos mutuamente excluyente y exhaustivo** ya que, de todos ellos, sólo uno debe ocurrir y no pueden ocurrir dos simultáneamente.

### Diagramas de árbol

Este diagrama permite indicar de manera sencilla el conjunto de posibles resultados en un experimento aleatorio siempre y cuando los resultados del experimento puedan obtenerse en diferentes fases sucesivas.

### Combinaciones, Variaciones y Permutaciones

#### Combinaciones

Se denominan combinaciones de  $m$  elementos tomados de  $n$  en  $n$  al número de subconjuntos diferentes de  $n$  elementos que se pueden formar con los  $m$  elementos del conjunto inicial.

#### Combinaciones con repetición

Si en los subconjuntos anteriores se pueden repetir los elementos

#### Variaciones

Se denominan variaciones de  $m$  elementos tomados de  $n$  en  $n$  a los distintos subconjuntos diferentes de  $n$  elementos que se pueden formar con los  $m$  elementos, influyendo el orden en el que se toman.

## PROBABILIDAD

Interesa ahora determinar la medida numérica de la posibilidad de que ocurra un suceso  $A$  cuando se realiza el experimento aleatorio. A esta medida se la denomina **Probabilidad del suceso  $A$ :  $P(A)$** . Para cada punto en el espacio muestral de un experimento es posible asignar un número llamado *probabilidad*.

Existen algunas definiciones de probabilidad, desde el punto de vista subjetivo y desde el punto de vista objetivo.

### **Probabilidad a priori (clásica)**

Esta definición de probabilidad fue una de las primeras que se presentaron (1900) y se atribuye a Laplace; también se conoce con el nombre de **probabilidad a priori** pues, para calcularla, es necesario conocer, antes de realizar el experimento aleatorio, el espacio muestral y el número de resultados o sucesos elementales que entran a formar parte del suceso.

Esta definición (la más antigua) se originó en los juegos de azar, y se basa en el sencillo supuesto de resultados **igualmente probables** de un experimento.

Sea un experimento aleatorio cuyo correspondiente espacio muestral  $E$  está formado por un número  $N$  finito de posibles resultados distintos y con la misma probabilidad de ocurrir  $\{e_1, e_2, \dots, e_n\}$

Si  $n_1$  resultados constituyen el subconjunto o suceso  $A_1$ ,  $n_2$  resultados constituyen el subconjunto o suceso  $A_2$  y, en general,  $n_k$  resultados constituyen el subconjunto o suceso  $A_k$  de tal forma que:

$$n_1 + n_2 + \dots + n_k = N$$

Luego la probabilidad de cualquier suceso  $A$  de este espacio muestral es igual al cociente entre el número de casos favorables que integran el suceso  $A$  y el número de casos posibles del espacio muestral  $E$ .

$$P(A) = \frac{\text{N}^\circ \text{ de casos favorables de } A}{\text{N}^\circ \text{ de casos posibles de } E} = \frac{n}{N} \quad (1.1)$$

La aplicación de la definición clásica de probabilidad puede presentar dificultades de aplicación cuando el espacio muestral es infinito o cuando los posibles resultados de un experimento no son equiprobables. Por ej: En un proceso de fabricación de piezas puede haber algunas defectuosas y si queremos determinar la probabilidad de que una pieza sea defectuosa no podemos utilizar la definición clásica pues necesitaríamos conocer previamente el resultado del proceso de fabricación; en el caso de determinar la probabilidad de un día de lluvia en un mes determinado se debería poder tener el registro de lo ocurrido hasta ese momento.

Para resolver estos casos, se hace una extensión de la definición de probabilidad, de manera que se pueda aplicar con menos restricciones, llegando así a la definición frecuentista de probabilidad.

## Probabilidad frecuencial o a posteriori

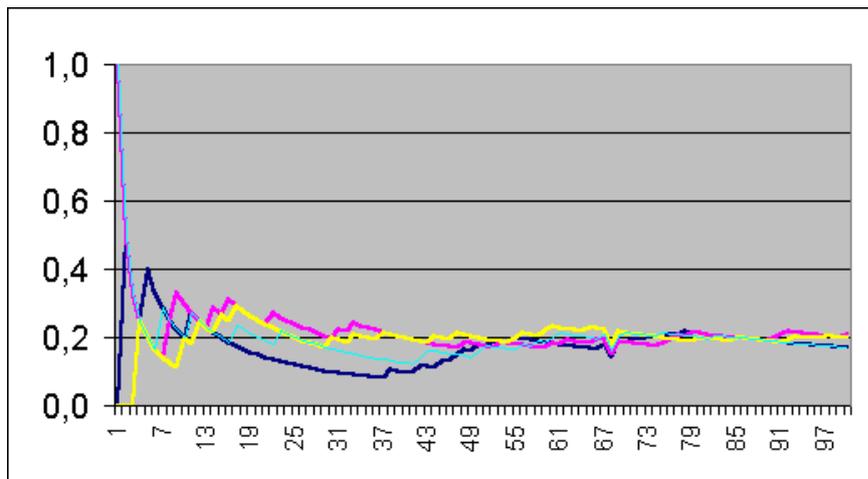
Si un experimento es repetido un número  $n$  de veces en condiciones iguales y existen  $n_1$  ( $n_1 \leq n$ ) resultados en los cuales se ha presentado el suceso que interesa,  $n_1/n$  es la frecuencia relativa del suceso, si el número de repeticiones se hace muy grande esa frecuencia relativa converge a la probabilidad del suceso.

$$P(B) = \lim_{n \rightarrow \infty} \frac{n_1}{n} \quad (1.2)$$

Es claro que en la práctica se debe tener un límite y en realidad se toma a la frecuencia relativa como la probabilidad del evento.

Esta definición frecuentista de la probabilidad se llama también **probabilidad a posteriori** ya que sólo podemos dar la probabilidad de un suceso después de repetir y observar un gran número de veces el experimento aleatorio correspondiente.

En la gráfica siguiente se observa la repetición de una serie de experimentos una determinada cantidad de veces y la estabilización de las frecuencias hacia un valor constante que es su probabilidad.



## Probabilidad subjetiva

Tanto la definición clásica como la frecuentista se basan en las repeticiones del experimento aleatorio; pero existen muchos experimentos que no se pueden repetir bajo las mismas condiciones y por tanto no puede aplicarse la interpretación objetiva de la probabilidad.

En esos casos es necesario acudir a un punto de vista alternativo, que no dependa de las repeticiones, sino que considere la probabilidad como un concepto **subjetivo** que exprese el grado de creencia o confianza individual sobre la posibilidad de que el suceso ocurra.

Se trata por tanto de un juicio personal o individual y es posible que diferentes observadores tengan distintos grados de creencia sobre los posibles resultados, igualmente válidos.

Se puede definir como la probabilidad asignada a un evento por parte de un individuo, basada en la evidencia que se tenga disponible. Esa evidencia puede presentarse en forma de frecuencia relativa de presentación de eventos pasados o puede tratarse simplemente de una creencia meditada.

Los tomadores de decisiones pueden hacer uso de cualquier evidencia que tengan a mano y mezclarlas con los sentimientos personales sobre la situación.

Como casi todas las decisiones sociales y administrativas de alto nivel se refieren a situaciones específicas y únicas, los responsables de tomar decisiones hacen un uso considerable de la probabilidad subjetiva.

### **Probabilidad axiomática**

La definición axiomática de la probabilidad es quizás la más simple de todas las definiciones y la menos controvertida ya que está basada en un conjunto de axiomas que establecen los requisitos mínimos para dar una definición de probabilidad.

La ventaja de esta definición es que permite un desarrollo riguroso y matemático de la probabilidad. Fue introducida por A. N. Kolmogorov y aceptada por estadísticos y matemáticos en general.

Se dice que un número **P** es la probabilidad de un evento si satisface los tres axiomas siguientes:

\***Axioma I:** La probabilidad de un evento es un número mayor o igual que cero y menor o igual que uno.

$$0 \leq P(A) \leq 1 \quad (1.3)$$

\***Axioma II:** La probabilidad del espacio muestral **E** es igual a uno.

$$P(E) = 1 \quad (1.4)$$

\***Axioma III:** La probabilidad de un evento el cual es la unión de dos eventos mutuamente excluyentes es la suma de sus probabilidades.

$$A \cap B = \emptyset \\ P(A \cup B) = P(A) + P(B)$$

Esto puede generalizarse como:

$$P(\cup_i A_i) = \sum_{i=1}^n P(A_i), \quad \text{con } A_i \cap A_j = \emptyset \quad \text{y } i \neq j \quad (1.5)$$

Se desprenden algunas consecuencias de estos axiomas que tiene mucha utilidad práctica:

a – Si  $\bar{A}$  es el complemento de  $A$  en el espacio muestral  $E$ , luego:

$$\bar{A} \cup A = E$$

$$\bar{A} \cap A = \emptyset$$

$$P(\bar{A} \cup A) = P(E) \quad \text{y} \quad P(E) = 1 \quad (1.6)$$

$$P(\bar{A}) + P(A) = 1$$

$$P(\bar{A}) = 1 - P(A)$$

b - La probabilidad del suceso imposible es igual a cero. El suceso contrario de  $E$  es un suceso imposible, por la consecuencia anterior:

$$P(\bar{E}) = 1 - P(E) = 1 - 1 = 0 \quad (1.7)$$

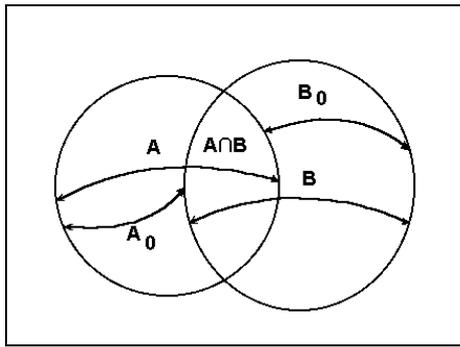
$$P(\emptyset) = 0$$

- Si para cualquier suceso  $A$  resulta que  $p(A)=0$  diremos que  $A$  es el suceso **nulo**, pero esto no implica que  $A = \emptyset$
- Si para cualquier suceso  $A$  resulta que  $p(A)=1$  diremos que  $A$  es el suceso **casi seguro**, pero esto no implica que  $A = E$

Ciertas relaciones entre las probabilidades de eventos se desprenden de las relaciones entre eventos y de los axiomas de probabilidad. Algunas conclusiones simples son evidentes, pero otras necesitan de algunas demostraciones.

## PROBABILIDAD TOTAL

**Teorema.** Si dos eventos  $A$  y  $B$  pertenecen al mismo espacio muestral, la probabilidad de que  $A$  o  $B$  o ambos ocurran *es la suma de sus probabilidades menos la probabilidad de su ocurrencia conjunta.*



$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (1.8)$$

Esta última expresión se justifica considerando dos sucesos no excluyentes A y B:

$$A = (A \cap B) \cup A_0$$

$$B = (A \cap B) \cup B_0$$

Sus probabilidades, por el Axioma III, son:

$$P(A) = P(A \cap B) + P(A_0)$$

$$P(B) = P(A \cap B) + P(B_0)$$

Luego, la unión de A y B será:

$$A \cup B = A_0 \cup (A \cap B) \cup B_0$$

Entonces, por el Axioma III:

$$P(A \cup B) = P(A_0) + P(A \cap B) + P(B_0)$$

De las expresiones anteriores, es posible obtener  $P(A_0)$  y  $P(B_0)$  para reemplazarlas en esta última:

$$P(A_0) = P(A) - P(A \cap B)$$

$$P(B_0) = P(B) - P(A \cap B)$$

$$P(A \cup B) = P(A) - P(A \cap B) + P(A \cap B) + P(B) - P(A \cap B)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

lo que demuestra la expresión correspondiente a 1.8

Esto se generaliza para **n** sucesos:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^n p(A_i) - \sum_{i<j} p(A_i \cap A_j) + \sum_{i<j<k} p(A_i \cap A_j \cap A_k) + \dots + (-1)^{n+1} p\left(\bigcap_{i=1}^n A_i\right)$$

### Árbol de Probabilidad

Es una herramienta que se utiliza para determinar todos los posibles resultados de experimentos aleatorios.

En el cálculo de la probabilidad se requiere conocer el número de elementos que forman parte del espacio muestral estos se pueden determinar con la construcción del diagrama de árbol.

Reglas para probabilidad utilizando diagramas de árbol:

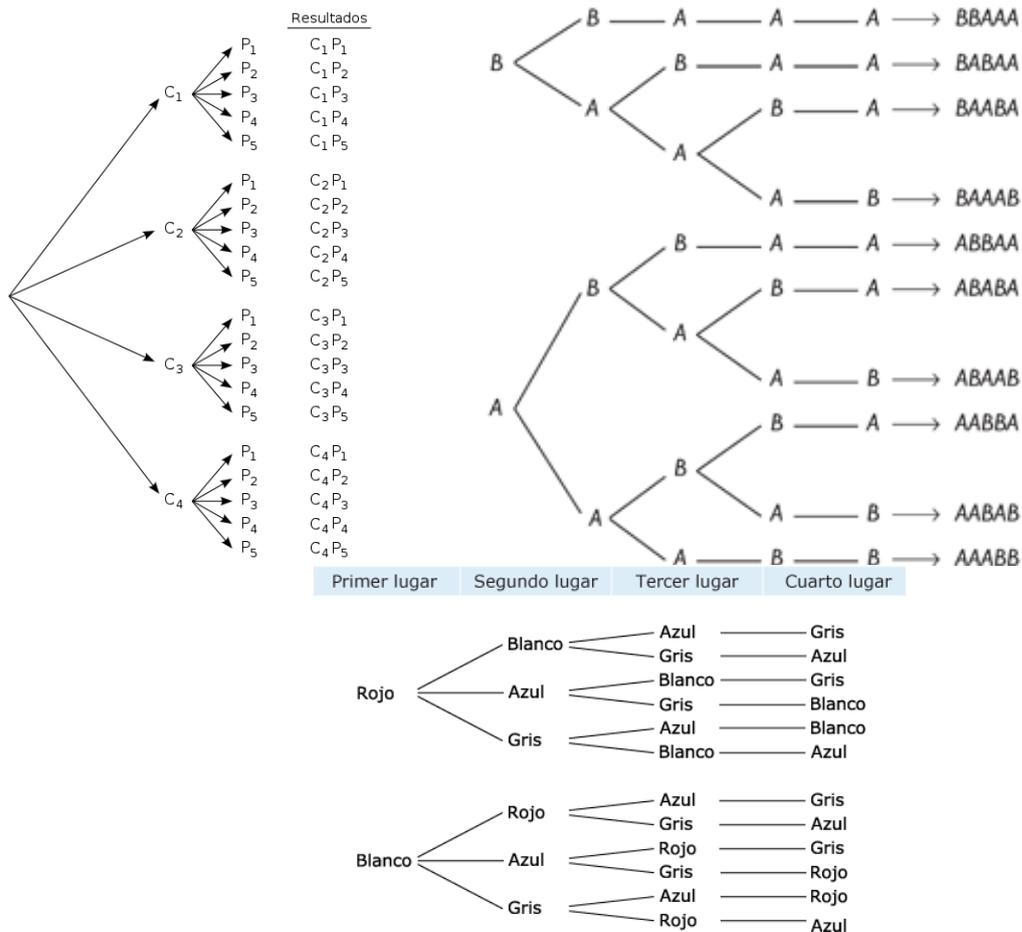
- 1.- la probabilidad de un evento es la multiplicación de las probabilidades de la rama.
- 2.- la probabilidad de varios eventos es la suma de las probabilidades de las ramas con casos favorables al evento.
- 3.- los eventos son mutuamente excluyentes en 2 o más ramas

Para la construcción de un **diagrama en árbol** se partirá poniendo una **rama** para cada una de las **posibilidades**, acompañada de su **probabilidad**.

En el **final** de cada **rama parcial** se constituye a su vez un **nudo** del cual parten nuevas **ramas**, según las **posibilidades** del siguiente paso, salvo si el nudo representa un posible final del experimento (**nudo final**).

Hay que tener en cuenta que la **suma de probabilidades** de las **ramas** de cada **nudo** debe dar **1**.

**Ejemplos:**



## PROBABILIDAD COMPUESTA Y CONDICIONAL

Hasta ahora hemos introducido el concepto de probabilidad considerando que la única información sobre el experimento era el espacio muestral. Sin embargo hay situaciones en las que se incorpora información suplementaria respecto de un suceso relacionado con el experimento aleatorio, cambiando su probabilidad de ocurrencia.

El hecho de introducir más información, como puede ser la ocurrencia de otro suceso, conduce a que determinados sucesos no pueden haber ocurrido, variando el espacio de resultados y cambiando sus probabilidades.

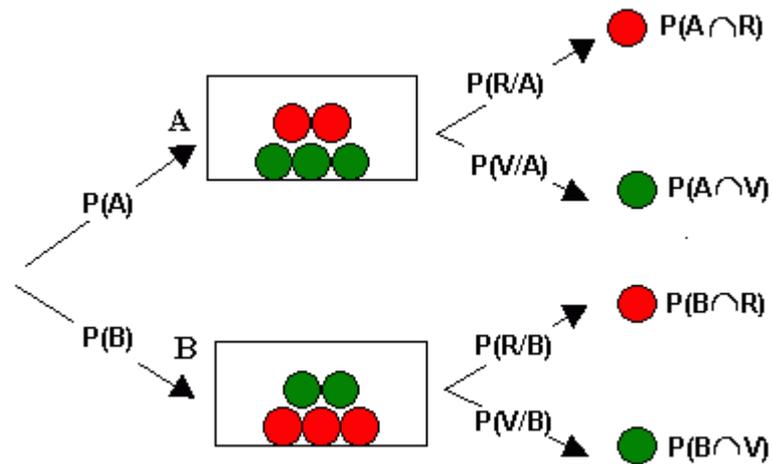
Consideremos ahora el siguiente experimento:

Dos cajas, **A** y **B**, contienen tubos con agua para la realización de análisis químicos; la caja **A**, contiene 3 tubos con agua potable (**V**) y 2 tubos con agua de río (**R**), la caja **B** contiene 2 tubos con agua potable y 3 tubos con agua de río.

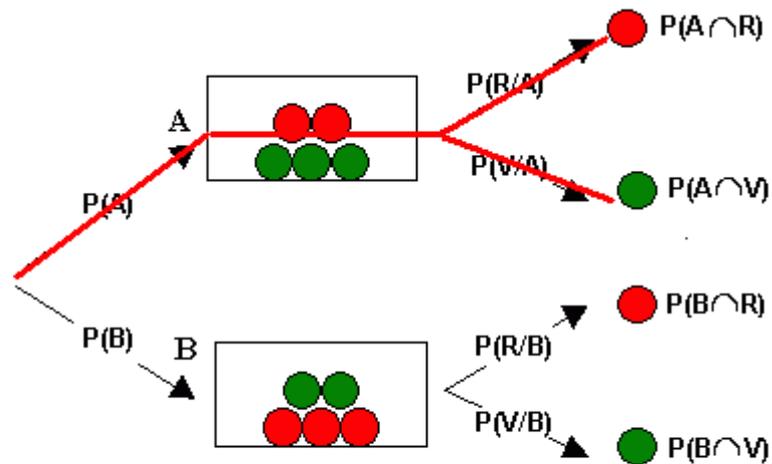
Se realiza el experimento en dos tiempos, primero se selecciona una caja por un procedimiento aleatorio y posteriormente de la caja elegida se extrae un tubo.

Para representar, de forma adecuada, este tipo de experimentos, se realiza un árbol de

probabilidades:



Si se sabe que ha ocurrido el suceso  $A$ , hay que volver a asignar probabilidades a los distintos caminos; todos los caminos que comienzan por el suceso  $B$ , tendrán probabilidad 0 y los que empiezan por el suceso  $A$  son los señalados en rojo:



$$P(A) = P(A \cap R) + P(A \cap V)$$

Por lo tanto en números sería:

$$P(R/A) = \frac{2}{10}$$

Esto conduce al siguiente resultado:

$$P(R/A) = \frac{P(A \cap R)}{P(A)} \Leftrightarrow P(A \cap R) = P(A) \times P(R/A) \quad (1.9)$$

### Definición 1. Probabilidad condicionada o condicional

De un suceso  $R$  sabiendo que ha ocurrido otro  $A$

$$P(R/A) = \frac{P(A \cap R)}{P(A)}$$

### Teorema 1. Regla del producto

De la definición anterior, se deduce:

$$P(A \cap R) = P(A) \times P(R/A)$$

### Sucesos dependientes

Dos sucesos son dependientes si el resultado de uno influye en el otro. Los sucesos  $A$  y  $B$  son dependientes si y sólo si  $P(A)$  es distinto de  $P(A/B)$  y  $P(B)$  es distinto de  $P(B/A)$

### Sucesos independientes

Dos sucesos son independientes si el resultado de uno no influye en el resultado del otro. Los sucesos  $A$  y  $B$  son independientes si y sólo si

$$P(A) = P(A/B) \text{ y } P(B) = P(B/A).$$

Diremos que dos sucesos  $A$  y  $B$  son **independientes** si se verifica una cualquiera de las siguientes condiciones equivalentes:

- $p(B/A) = p(B)$  si  $P(A) > 0$
- $p(A/B) = p(A)$  si  $P(B) > 0$
- $p(A \cap B) = p(A)p(B)$

Podemos decir por lo tanto que si el suceso  $B$  es independiente del suceso  $A$ , entonces el suceso  $A$  también es independiente del suceso  $B$ , lo que equivale a decir que ambos sucesos son **mutuamente independientes**.

La independencia de sucesos puede extenderse a más de dos sucesos:

$$p(A \cap B \cap C) = p(A)p(B)p(C)$$

Además, se cumple el siguiente teorema: Si  $A$  y  $B$  son dos sucesos independientes, entonces también lo son los sucesos  $\bar{A}$  y  $B$ ;  $A$  y  $\bar{B}$ ;  $\bar{A}$  y  $\bar{B}$

### Sucesos independientes y excluyentes

Es importante tener clara la diferencia entre sucesos independientes y excluyentes. El primer caso se analiza cuando se considera la ocurrencia conjunta de ambos sucesos; en cambio, se habla de sucesos excluyentes cuando se considera la unión de sucesos, o sea cuando se quiere determinar la probabilidad de que por lo menos uno de los sucesos ocurra.

-Si dos sucesos son excluyentes y si ninguno tiene probabilidad igual a cero, *luego deben ser estadísticamente dependientes*. Sin embargo, el inverso no puede asegurarse directamente, o sea, si dos sucesos son dependientes que sean también mutuamente excluyentes:

$$P(A \cup B) = P(A) + P(B), \text{ siendo } P(A) \text{ y } P(B) \neq 0$$

$$A \cap B = \emptyset$$

$$P(A \cap B) = 0$$

Dos sucesos son dependientes si:

$$P(A \cap B) \neq P(A)P(B)$$

$$0 \neq P(A)P(B), \text{ ya que } P(A) \text{ y } P(B) \neq 0$$

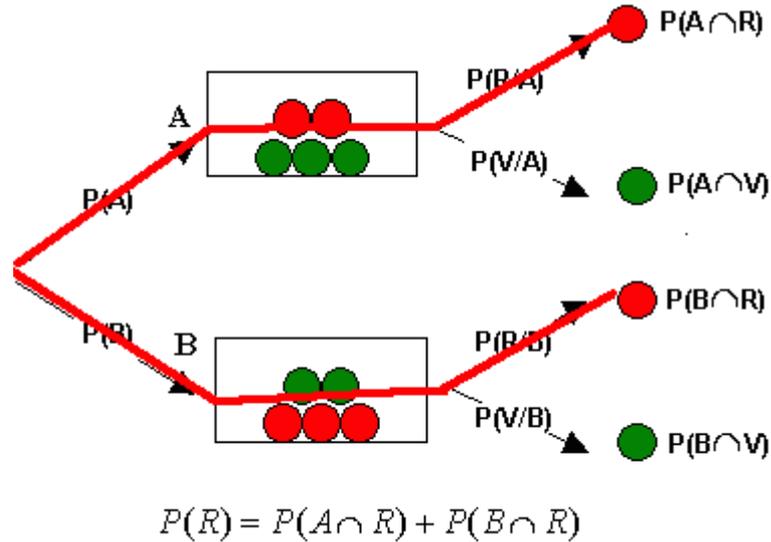
-Si dos eventos son independientes y si sus probabilidades son distintas de cero entonces puede decirse que *no pueden ser excluyentes* porque para dos sucesos independientes la probabilidad de que los dos sucedan es el producto de sus probabilidades:

$P(A \cap B) = P(A) P(B)$ , y esto será distinto de cero; luego la probabilidad de la unión será:  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ , expresión que corresponde a sucesos no excluyentes.

### TEOREMA DE BAYES (Probabilidad a posteriori)

Se considera de nuevo el experimento de las cajas A y B, que contienen tubos verdes y rojos:

Si se sabe que ha salido un tubo rojo, los caminos posibles en el árbol de probabilidades, quedan reducidos a dos, los señalados en rojo en el gráfico anterior; tenemos que reasignar probabilidades, todos los caminos que terminan en tubo verde, deberán tener probabilidad 0. ¿Cómo asignamos probabilidades a los caminos que conducen a tubo rojo?



$$P(A/R) = \frac{P(A \cap R)}{P(A \cap R) + P(B \cap R)} = \frac{P(A) \cdot P(R/A)}{P(A) \cdot P(R/A) + P(B) \cdot P(R/B)} \quad (1.10)$$

**Generalizando:** Supóngase un conjunto de eventos mutuamente excluyentes  $(A_1, A_2, \dots, A_n)$  con probabilidades mayores que cero, respecto a un experimento aleatorio y que pueden considerarse como un conjunto de hipótesis acerca del suceso que interesa (R) y pertenecientes al espacio muestral. Si el experimento se ha realizado y se conoce que el suceso R ha ocurrido, la probabilidad a posteriori, de que se realice la hipótesis  $A_i$ , a condición de que R haya ocurrido, se obtiene a través de la fórmula de **Bayes**.

$$P(A_i/R) = \frac{P(A_i) \cdot P(R/A_i)}{\sum_{i=1}^n P(A_i) \cdot P(R/A_i)} \quad \text{con } i = 1, 2, \dots, n \quad (1.11)$$

De esta manera la ocurrencia de R condiciona la información previa  $P(A_i)$ . Estas probabilidades  $P(A_i)$  son llamadas probabilidades a priori, ya que son las probabilidades de las causas  $A_i$  antes de observar R. De la misma manera  $P(A_i/R)$  es llamada probabilidad a posteriori ya que se refiere al estado de conocimiento de  $A_i$  después de la ocurrencia de R. Se ve que la fórmula de Bayes permite con información experimental nueva, actualizar la que existe relacionada a un evento.

### Ejemplo:

Se planea mejorar el manejo de los datos hidrometeorológicos de una cuenca dada mediante la utilización del sistema de radares ubicados en tres sitios de la misma. Los sistemas de radar en cada sitio  $A_1$ ,  $A_2$ ,  $A_3$ , se ponen a funcionar el 30%, 20% y 50% del tiempo respectivamente. Si un técnico que tiene a su cargo el sistema recibe datos de tipos aceptable y deficientes de los tres lugares: del sitio 1 el 3% defectuosos, del sitio 2 5% defectuosos, del sitio 3 4% defectuosos; ¿cuál es la probabilidad de que el técnico registre un conjunto de datos deficientes? y cuál la probabilidad de que habiendo recibido un conjunto de datos deficientes provengan del radar ubicado en el sitio 2?

Solución:

$H_1$ : funciona y envía el radar en el sitio  $A_1$                        $P(A_1) = 0.3$

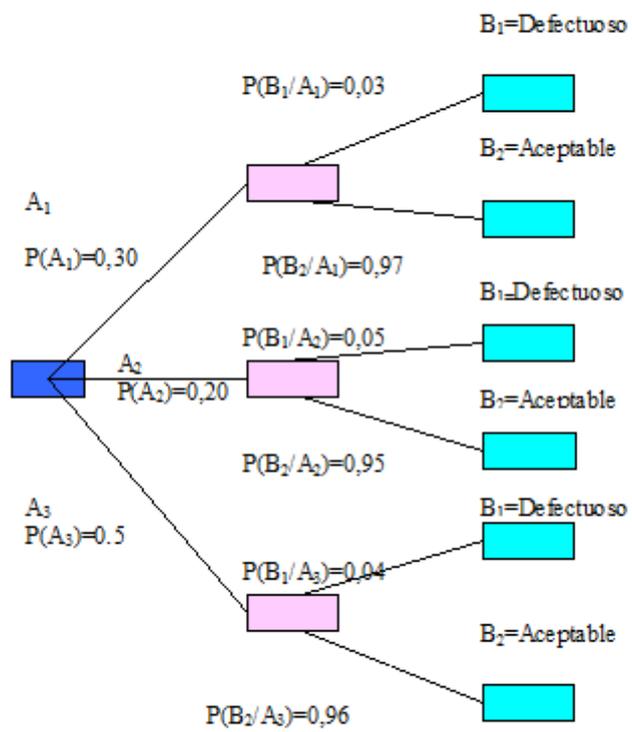
$H_2$ : funciona y envía el radar en el sitio  $L_2$                        $P(A_2) = 0.2$

$H_3$ : funciona y envía el radar en el sitio  $A_3$                        $P(A_3) = 0.5$

$E$ : que el técnico registre un conjunto de datos defectuosos

$$P(E) = P(A_1) \cdot P(E/A_1) + P(A_2) \cdot P(E/A_2) + P(A_3) \cdot P(E/A_3) = 0.3 \cdot 0.03 + 0.2 \cdot 0.05 + 0.5 \cdot 0.04 = 0.039$$

$$P(A_2/E) = \frac{P(A_2) \cdot P(E/A_2)}{P(E)} = \frac{0.2 \cdot 0.05}{0.039} = 0.256$$





Universidad Nacional del Litoral  
**Facultad de Ingeniería y Ciencias Hídricas**

## **ESTADÍSTICA**

**Ingenierías: Recursos Hídricos-  
Ambiental-Agrimensura**

---

TEORÍA

*Mg.Ing. Susana Vanlesberg*  
Profesor Titular

**UNIDAD 2**  
**VARIABLES ALEATORIAS**

## 2.-VARIABLES ALEATORIAS Y DISTRIBUCIÓN DE PROBABILIDADES

En el capítulo I al hablar de sucesos elementales se definieron casi exclusivamente en forma cualitativa: buenos y malos; defectuosos y no defectuosos; llueve y no llueve, etc.; pero muchas veces es necesario cuantificar resultados. Es más, la mayor parte de los problemas que se plantean en ingeniería tienen que ver con cantidades medibles. Esta cuantificación se realiza a través de una variable que será de tipo *aleatoria*.

Formalmente definida *variable aleatoria* es aquella función que asocia a cada suceso elemental del espacio muestral un número perteneciente a los reales.

Es decir, una variable aleatoria es una variable cuyo valor numérico está determinado por el resultado del experimento aleatorio. Al realizar el experimento aleatorio la variable toma un solo valor.

La variable aleatoria (v.a.) será notada con letras en mayúscula X, Y, .....y con letras en minúscula x, y, .....sus valores.

La v.a. puede tomar una cantidad numerable o no numerable de valores, dando lugar a dos tipos de v.a.: discretas y continuas.

### Clasificación

Si el fenómeno estudiado u observado es tal que puede ser cuantificado por valores contables o enteros, entonces la variable aleatoria se dice **discreta**. El conjunto de los valores posibles de esta variable puede ser finito o infinito numerable por ejemplo contar el número de inundaciones en los últimos 10 años en la ciudad de Santa Fe, números de eventos de contaminación producidos por los efluentes de una empresa, cantidad de teodolitos que se llevan a un trabajo de campo.

Si el fenómeno puede ser cuantificado a través de cualquier número real, la variable se dirá **continua**. Por ejemplo monto de lluvia precipitada en cierto día en Santa Fe; descarga diaria de un canal, contenido de un contaminante en un efluente, medida del lado de una poligonal

### 2.1- DISTRIBUCIÓN DE PROBABILIDADES

Si se considera la variable aleatoria discreta "cantidad de días de lluvia en un año" se verifica que hay cierta cantidad de resultados posibles: pueden haber 20, 30, 150 etc. También puede asegurarse que hay una cierta probabilidad que sean 30 días con lluvia, que sean 150 etc.

Existe una asignación de probabilidades a cada uno de los posibles valores que puede tomar una variable aleatoria. Esa correspondencia entre los valores de la variable aleatoria y su probabilidad de ocurrencia es lo que se denomina

## Distribución de Probabilidades

El comportamiento de una variable aleatoria es descrito por sus leyes de probabilidad, las cuales pueden ser caracterizadas de distintas formas:

- a- la forma más común es a través de la distribución de probabilidades, el caso más sencillo es a través de una lista de los valores que la variable aleatoria puede tomar y sus respectivas probabilidades (sólo posible en el caso de variable aleatoria discreta);
- b- a través de gráficos
- c- a través de modelos matemáticos que representen la ley.

### 2.1.1 - Variable Discreta

#### Función de Cuantía

Esta función es la forma matemática de expresar la correspondencia entre los valores de la variable y sus probabilidades:

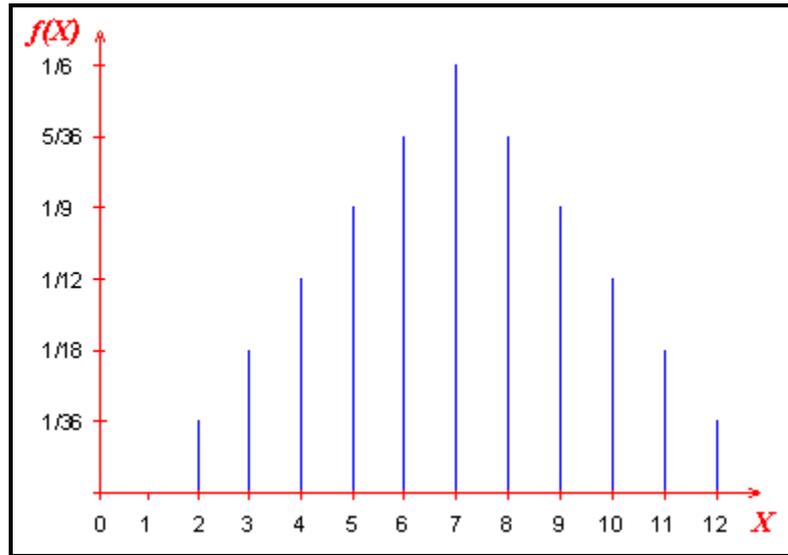
$$f(x) = P(X = x) \quad (1)$$

Para ser una **función de probabilidad**, deberá cumplir con los siguientes requisitos:

a-  $f(x_i)$  debe tener un valor numérico para todos los posibles valores de la variable aleatoria.

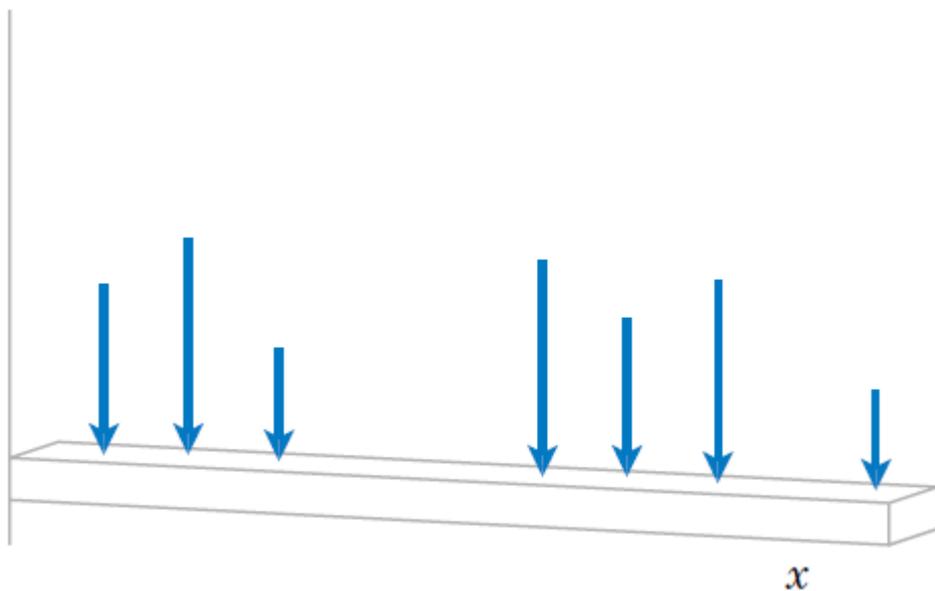
$$b- 0 \leq f(x_i) \leq 1 \quad i=1,2,\dots,r$$

$$c- \sum_{i=1}^r f(x_i) = 1$$



**Fig. N° 1- Función de cuantía**

Puede hacerse una semejanza con mecánica considerando una masa total unitaria distribuida sobre el eje de las  $x$ , con la masa ubicada en cada uno de los puntos y proporcional en altura a la probabilidad de que la variable tome ese valor.



**Función de Distribución**

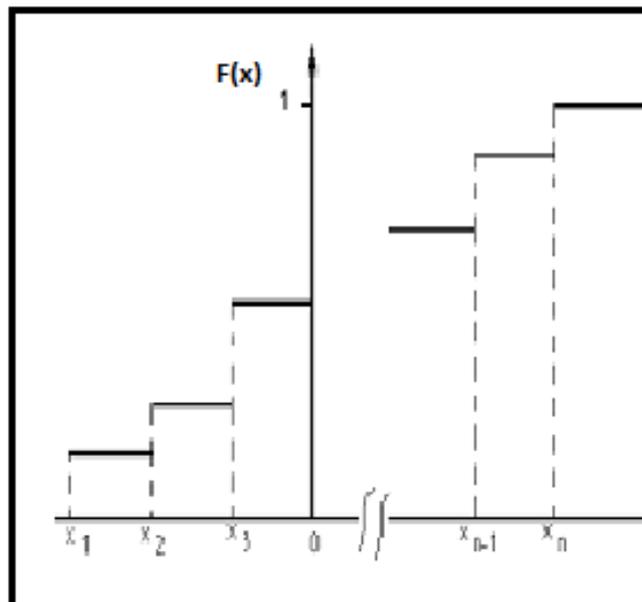
Otra forma de describir la distribución de probabilidades de una variable aleatoria es a través de una función denominada **función acumulativa** o **función de distribución**. Esta función  $F(x)$  es simplemente la probabilidad que la variable aleatoria tome valores menores o iguales que un valor determinado:

$$F(x) = P(X \leq x) \quad (2)$$

Para el caso de variable aleatoria discreta esta función es la suma de los valores de la función masa de probabilidad evaluada en todos aquellos valores menores o iguales que  $x$ , que la variable puede tomar:

$$F(x) = \sum_{\forall x_i \leq x} f(x_i) \quad (3)$$

La gráfica de esta función es de tipo escalonada, ya que experimenta saltos en los distintos  $x_i$ , iguales a la probabilidad de que la variable aleatoria tome esos valores. Entre los valores, la probabilidad permanece constante:



**Fig. N° 2- Función de distribución**

La función de distribución de una v.a. discreta es una función que verifica:

- $0 \leq F(x) \leq 1, \quad \forall x$
- Es no decreciente, es decir, si  $x_i \leq x_j$ , entonces  $F(x_i) \leq F(x_j)$

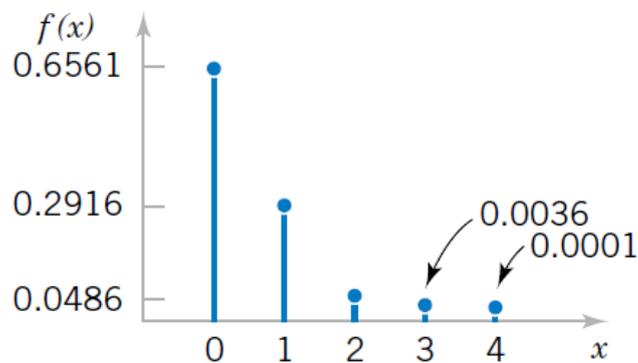
Se puede decir por tanto que una v.a.  $X$  discreta, está caracterizada por su función de probabilidad o distribución de probabilidad  $P(x)$  y también por su función de distribución  $F(x)$ .

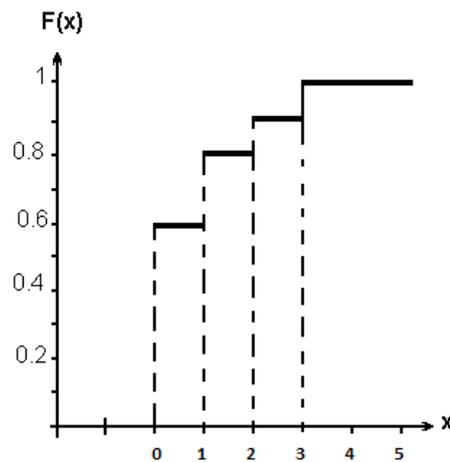
**Ejemplo:**

Considérese la variable aleatoria  $X$ : N° de barcos de carga que llegan a un puerto en una semana determinada del mes de mayo. Las probabilidades para los distintos casos son:

$$p(x) = \begin{cases} 0.6561 & x = 0 \\ 0.2916 & x = 1 \\ 0.0486 & x = 2 \\ 0.0036 & x = 3 \\ 0.0001 & x = 4 \\ 0 & x = 5, 6, \dots \end{cases}$$

$$F(x) = P(X \leq x) = \begin{cases} 0, & x < 0 \\ 0.6561, & 0 \leq x < 1 \\ 0.9477, & 1 \leq x < 2 \\ 0.9963, & 2 \leq x < 3 \\ 0.9999, & 3 \leq x < 4 \\ 1, & x \geq 4 \end{cases}$$





## 2.1.2.- Variable aleatoria continua

### Función de densidad

La v.a. de tipo continuo se tratará de forma diferente a como se ha visto en el caso de v.a. discreta, ya que en el caso continuo no es posible asignar una probabilidad a cada uno de los infinitos posibles valores de la v.a. y que estas probabilidades sumen uno (como en el caso discreto), teniendo por tanto que utilizar una aproximación diferente para llegar a obtener la distribución de probabilidad de una v.a. continua.

El problema de especificar la distribución de probabilidades y de esta manera la ley para una variable aleatoria continua es sencillo de derivar.

El eje x se divide en un gran número de intervalos infinitesimales, cada uno de longitud  $dx$ , es posible definir una función  $f(x)$ , tal que la probabilidad de que x esté en un intervalo  $(x, x+dx)$  es  $f(x)dx$ . Esta función se denomina *función de densidad de probabilidad* o, simplemente, *función de densidad*.

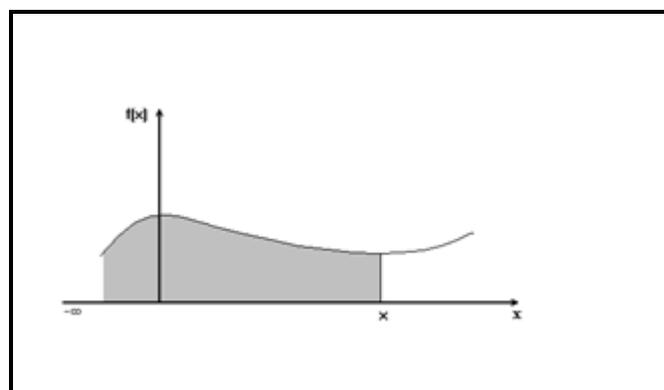
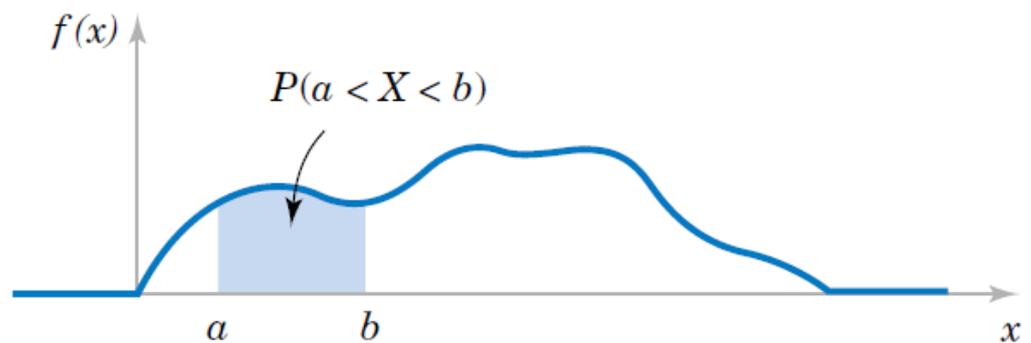


Fig. N° 3- Función de densidad

Ya que ocurrencias en distintos intervalos son eventos mutuamente excluyentes, se deduce que la probabilidad de que una variable aleatoria tome valores en un intervalo de longitud finita, es la suma de probabilidades en los intervalos infinitesimales, o sea, es  $\int f(x)dx$  sobre el intervalo de interés.

De esta manera, el área bajo la curva de esta función en un intervalo, representa la probabilidad de que la variable tome valores en el intervalo referido:

$$P(a \leq x \leq b) = \int_a^b f(x)dx \quad (4)$$



La probabilidad de que una variable aleatoria continua tome un valor específico es 0, ya que la longitud del intervalo  $dx$  desaparecería:

$$P(X = x_1) = \int_{x_1}^{x_1} f(x)dx = 0$$

Como consecuencia de esto, las siguientes expresiones son equivalentes:

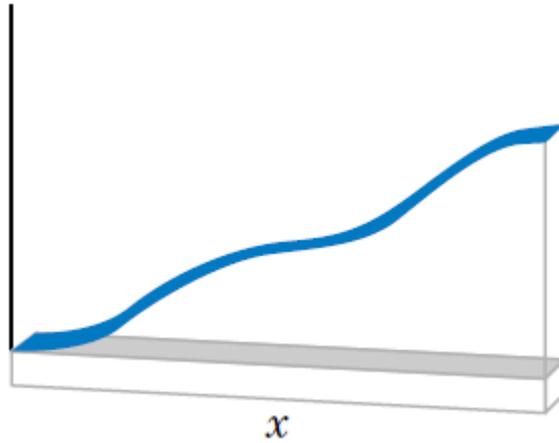
$$P(a \leq X \leq b) \quad P(a < X < b) \quad P(a \leq X < b) \quad P(a < X \leq b)$$

El valor de  $f(x)$  en sí mismo no es una probabilidad, sino solamente la medida de la densidad de probabilidad en un intervalo. Debe cumplir con las condiciones:

$$f(x) \geq 0$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Puede hacerse una analogía con mecánica, diciendo que se tiene una masa total unitaria distribuida en forma continua en el intervalo en el que toma valores la variable.

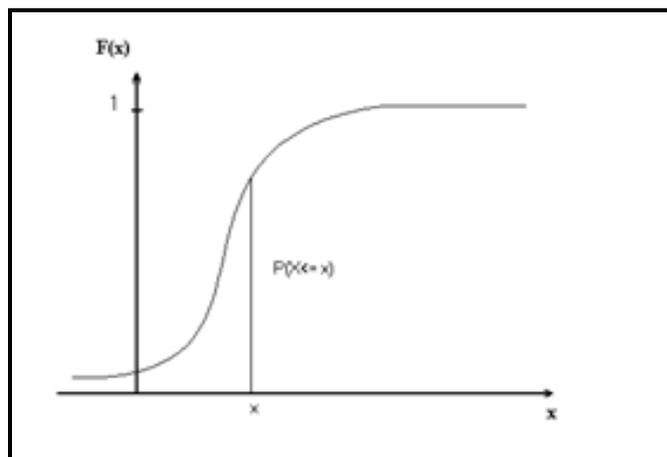


### **Función de Distribución**

Se define la función de distribución acumulativa de la v.a. X, F(x), como la probabilidad de que la v.a. continua X tome valores menores o iguales a x, es decir:

$$F(x) = P(X \leq x) \quad (5)$$

$$F(x) = \int_{-\infty}^x f(u) du \quad (6)$$



**Fig. N° 4- Función de distribución**

Consideramos una masa unidad distribuida sobre el intervalo  $(-\infty, +\infty)$ , y la función de distribución  $F(x)$  para cada valor  $x$  de la v.a. da la cantidad de masa que hay en el intervalo  $(-\infty, x]$ , es decir, la masa que hay en el punto  $x$  y a la izquierda de  $x$ , aunque dará igual considerar el intervalo  $(-\infty, x)$ .

**Propiedades de la Función de Distribución:**

$$F(-\infty) = 0, \quad F(x) = \lim_{x \rightarrow -\infty} \int_{-\infty}^x f(u) du = 0$$

$$F(\infty) = 1, \quad F(x) = \lim_{x \rightarrow \infty} \int_{-\infty}^x f(u) du = 1$$

Es una función monótona creciente:

Si  $x < x+m$  entonces  $F(x) \leq F(x+m)$ .

$$F(x+m) = F(x) + P(x < X \leq x+m)$$

Esto permite obtener la probabilidad en un intervalo:

$$F(x+m) - F(x) = P(x < X \leq x+m)$$

$$\boxed{P(a < X < b) = F(b) - F(a)}$$

Puede establecerse la relación entre ambas funciones, considerando un  $\Delta x$  arbitrariamente pequeño y evaluar la función acumulativa en los extremos de ese intervalo:  $x+\Delta x$  y  $x$ , haciendo tender a cero la amplitud  $\Delta x$ :

$$\lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}$$

$$\lim_{\Delta x \rightarrow 0} \frac{P(X \leq x + \Delta x) - P(X \leq x)}{\Delta x}$$

$$F'(x) = f(x) = \frac{P(x < X \leq x + \Delta x)}{\Delta x}$$

luego :

$$F(x) = \int_{-\infty}^x f(t) dt \quad (7)$$

## 2.2- Variables aleatorias bidimensionales

En determinadas ocasiones hay que trabajar en espacios de más de una dimensión, estableciendo aplicaciones que transforman los sucesos elementales del experimento aleatorio en puntos del espacio n-dimensional ( $\mathbb{R}^n$ ), estas aplicaciones se hacen utilizando v.a. bidimensionales o n-dimensionales.

En muchas ocasiones puede interesar estudiar conjuntamente dos características del fenómeno aleatorio, es decir, estudiar el comportamiento conjunto de dos v.a. para intentar explicar la posible relación entre ellas. Para poder estudiar conjuntamente las dos v.a., es necesario conocer la distribución de probabilidad conjunta.

Por ejemplo observar en forma simultánea cantidad de evaporación y promedio de temperatura en un día y lugar determinados.

Se puede definir variable aleatoria bidimensional como el par de números que expresa el resultado de un experimento combinado y esto puede ser representado en el plano x y; el recorrido o rango será ahora un subconjunto del plano x y.

También corresponde hacer la distinción entre variable bidimensional discreta y bidimensional continua.

## Distribución de probabilidades

### Variable aleatoria bidimensional discreta

#### Función masa de probabilidades conjuntas

Esta función es la correspondencia entre cada par de puntos y su probabilidad de ocurrencia. Analíticamente se expresa:

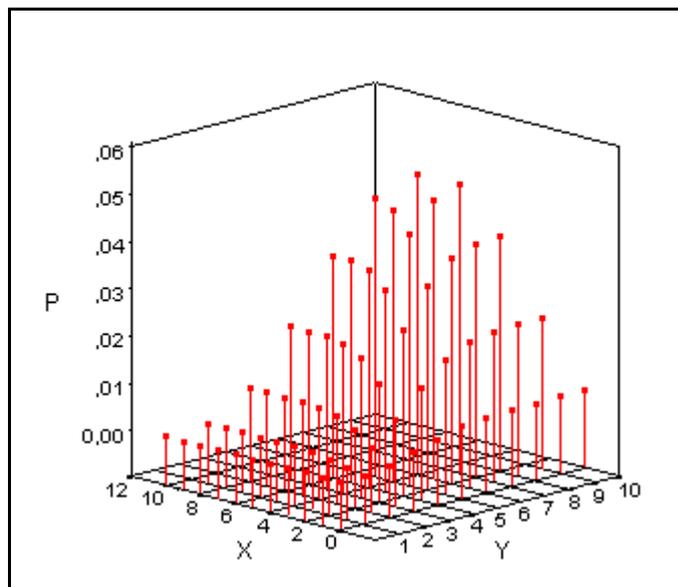
$$f(x, y) = P(X = x; Y = y) \quad (8)$$

ya que es una probabilidad debe cumplir con las siguientes condiciones :

$$f(x, y) \geq 0$$

$$\sum_{\forall x_i} \sum_{\forall y_j} f(x_i, y_j) = 1, \quad \text{con } i = 0, 1, \dots, m \quad j = 0, 1, \dots, n$$

Esto puede representarse en forma tridimensional. En el eje perpendicular al plano x y se grafican las probabilidades conjuntas.



**Fig. N° 5- Función de cuantía conjunta**

Esta distribución conjunta puede presentarse también a través de una tabla a doble entrada, considerando cada valor de cada variable en los extremos de la tabla y en el cuerpo de ella las probabilidades de presentación conjunta (para cada par de valores). Esto sólo puede hacerse si la variable aleatoria bidimensional es discreta. La función acumulativa conjunta se define a partir de la función masa:

$$F(x, y) = P(X \leq x; Y \leq y) = \sum_{\forall X_i \leq x} \sum_{\forall Y_j \leq y} f(x_i, y_j), i = 0, 1, \dots, m \quad j = 0, 1, \dots, n \quad (9)$$

<b>X \ Y</b>	<b>y1</b>	<b>y2</b>	<b>.....</b>	<b>Función marginal de X</b>
<b>x1</b>				
<b>x2</b>		<b>f(x2;y2)</b>		
<b>.....</b>				
<b>Función marginal de Y</b>				

## Funciones marginales

### Funciones masa de probabilidad marginales

El comportamiento de una variable en particular sin considerar la otra, se describe a través de las funciones marginales. Se las obtiene a partir de la función masa de probabilidades conjuntas. En forma analítica se expresa a través de:

$$f(x) = P(X = x) = \sum_{\forall y_j} f(x, y_j) \quad \text{con } j = 0, 1, \dots, n \quad (10)$$

$$f(y) = P(Y = y) = \sum_{\forall x_i} f(x_i, y) \quad \text{con } i = 0, 1, \dots, m$$

### Funciones acumulativas marginales

$$F(x) = P(X \leq x) = \sum_{\forall x_i \leq x} f(x_i) = \sum_{\forall x_i \leq x} \sum_{y_j=0}^n f(x_i, y_j) \quad (11)$$

$$F(y) = P(Y \leq y) = \sum_{\forall y_j \leq y} f(y_j) = \sum_{x_i=0}^m \sum_{\forall y_j \leq y} f(x_i, y_j)$$

Por ejemplo:

$$F(y_2) = P(X \leq x_m; Y \leq y_2) = \sum_{\forall x_i} \sum_{\forall Y \leq y_2} f(x_i, y_j)$$

## Funciones condicionales

Pueden obtenerse a partir de la distribución conjunta, es aquella que se deriva al evaluar las probabilidades de presentación de una variable conociendo que la otra variable ha ocurrido con un valor particular.

Para un valor particular de Y,  $Y=y_0$  la función de cuantía de x condicionada a Y está dada por :

$$f(x/y) = P(X = x/Y = y_0) = \frac{P(X = x \wedge Y = y)}{P(Y = y_0)}$$

$$f(x/y) = \frac{f(x,y)}{\sum_{\forall x_i} f(x_i, y)} = \frac{f(x,y)}{f(y)}, \text{ con } f(y) \neq 0 \quad (12)$$

Esta función de cuantía condicional, por ser una función de probabilidad debe cumplir con las siguientes condiciones:

$$a - 0 \leq f(x/y) \leq 1$$

$$b - \sum_{\forall x_i} f(x/y) = 1$$

La distribución condicional de Y dado un valor particular de X se define de forma similar:

$$f(y/x) = \frac{f(x,y)}{f(x)}, \text{ con } f(x) \neq 0 \quad (13)$$

## Variable aleatoria bidimensional continua

Las funciones asociadas con variables aleatorias bidimensionales continuas son análogas a las del caso discreto, pero la función de cuantía será reemplazada por la función de densidad conjunta, siendo  $f(x,y) \Delta x \Delta y$  la probabilidad de que la variable  $(x,y)$  caiga en el intervalo  $x, x+\Delta x$  ;  $y, y+\Delta y$ .

La probabilidad de ocurrencia conjunta de  $x$  e  $y$  en alguna región del espacio muestral se determina por integrar la función de densidad conjunta en esta región.

Esto es igual al volumen debajo de la función de densidad conjunta  $f(x,y)$ , sobre la región limitada por  $x_1, x_2$  e  $y_1, y_2$ .

$$P(x_1 \leq X \leq x_2; y_1 \leq Y \leq y_2) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(x,y) dx dy \quad (14)$$

La función de densidad conjunta debe cumplir con las siguientes condiciones:

$$f(x,y) \geq 0$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx dy = 1$$

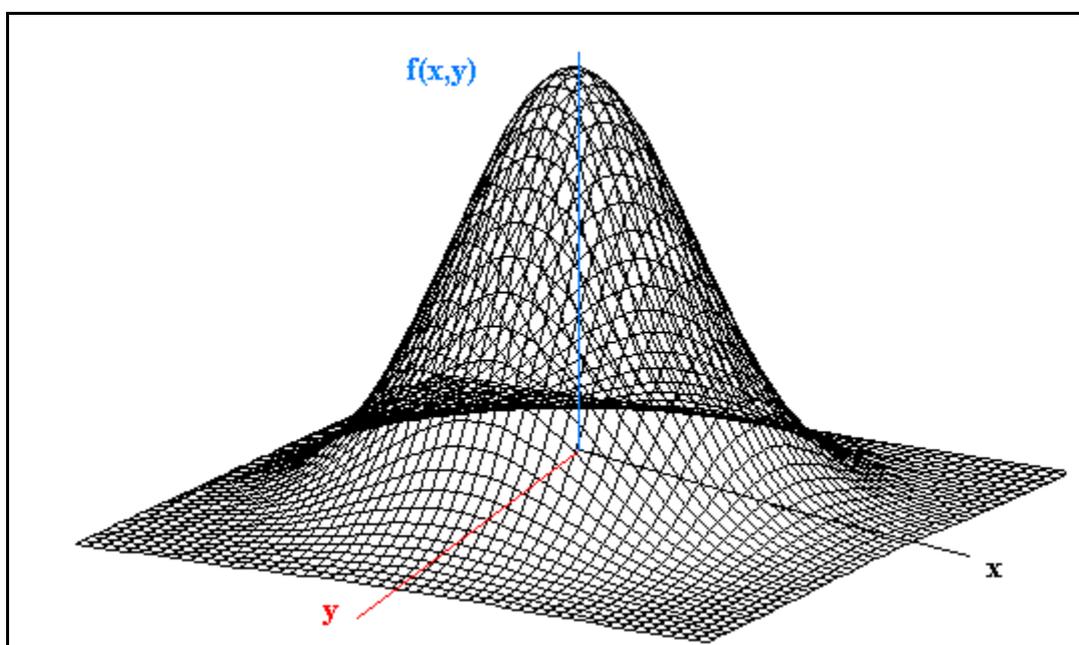


Fig. N° 6- Función de densidad conjunta

La función de distribución conjunta se define como:

$$F(x,y) = P(-\infty \leq X \leq x; -\infty \leq Y \leq y)$$

$$F(x,y) = P(X \leq x; Y \leq y) \quad (15)$$

$$F(x,y) = \int_{-\infty}^x \int_{-\infty}^y f(t,l) dt dl$$

### Propiedades

Son similares al caso unidimensional:

$$F(-\infty; -\infty) = 0$$

$$F(\infty; \infty) = 1$$

$$F(-\infty; y) = 0$$

$$F(x; -\infty) = 0$$

A partir de esta función, se puede obtener la de densidad conjunta por derivar parcialmente:

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y) \quad (16)$$

### Funciones de densidad marginales

Como en el caso discreto, para estudiar el comportamiento de una de las variables en particular se lo hace a través de las funciones marginales:

$$f(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

$$f(y) = \int_{-\infty}^{\infty} f(x,y)dx \quad (17)$$

### Funciones de distribución marginales

$$F(x) = P(X \leq x) = \int_{-\infty}^x \int_{-\infty}^{\infty} f(x,y)dydx = \int_{-\infty}^x f(t)dt$$

(18)

$$F(x) = F_{x,y}(x, \infty)$$

se deduce que:

$$f(x) = \frac{\partial}{\partial x} [F(x, \infty)] = \frac{\partial}{\partial x} [F(x)]$$

de forma similar:

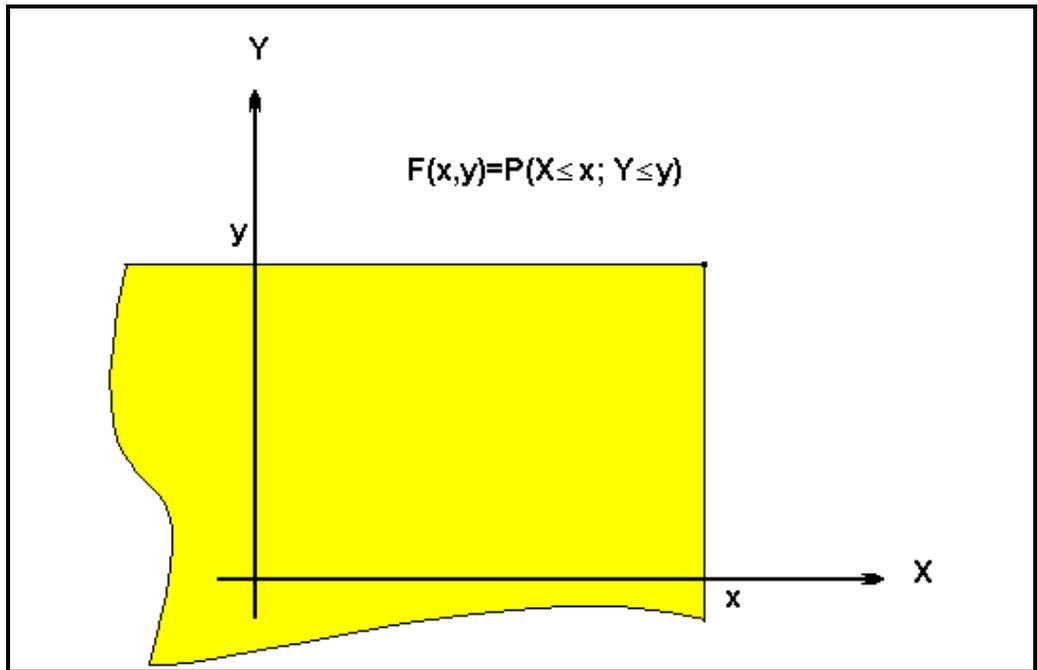
$$F(y) = P(Y \leq y) = \int_{-\infty}^{\infty} \int_{-\infty}^y f(x,y)dx dy = \int_{-\infty}^y f(s)ds$$

(19)

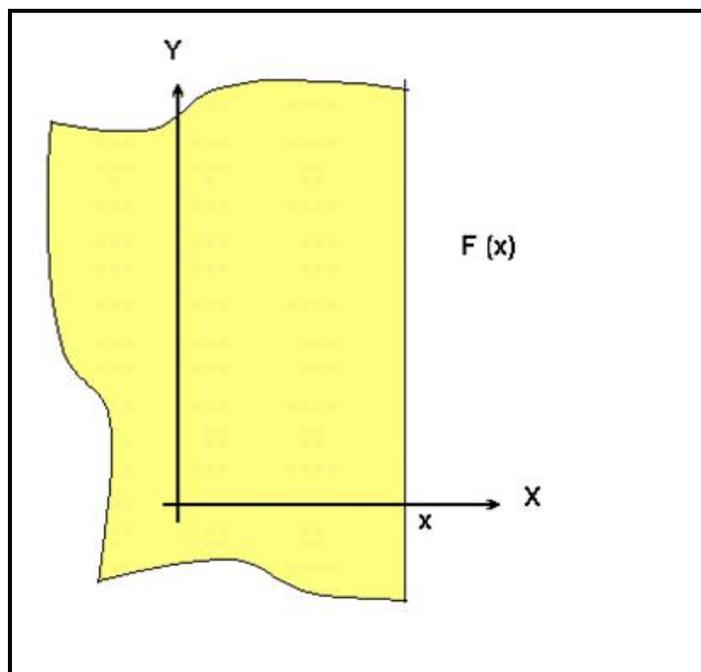
$$F(y) = F_{x,y}(\infty, y)$$

$$f(y) = \frac{\partial}{\partial y} F(\infty, y) = \frac{\partial}{\partial y} [F(y)]$$

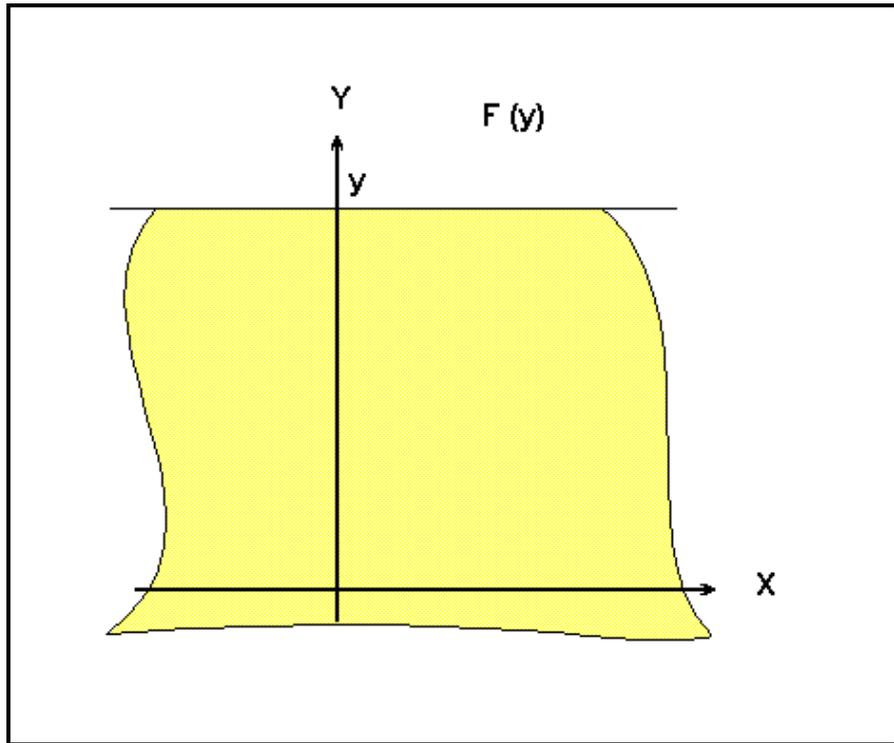
Gráficamente todas las funciones:



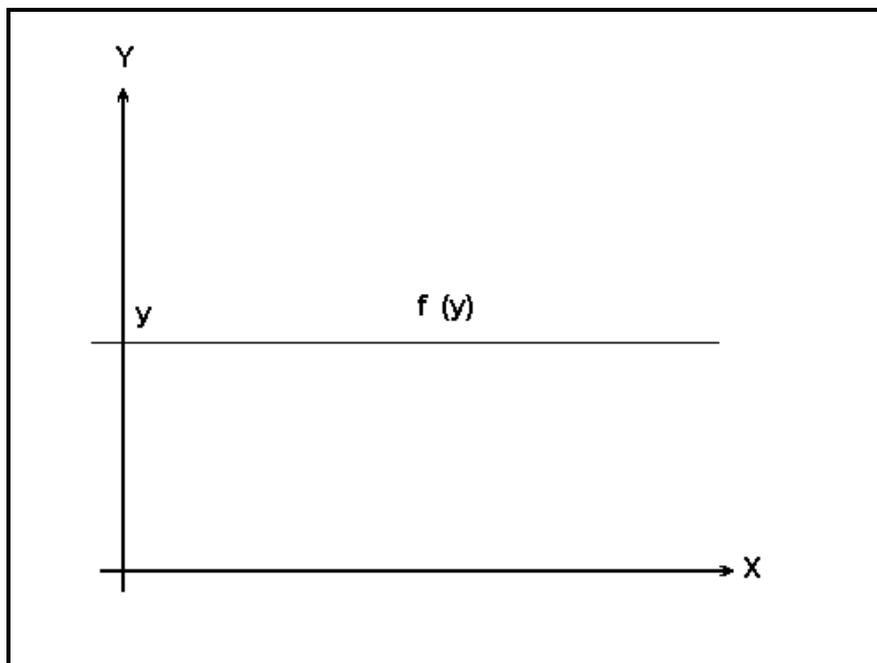
**Fig. N° 7- Función de distribución conjunta**



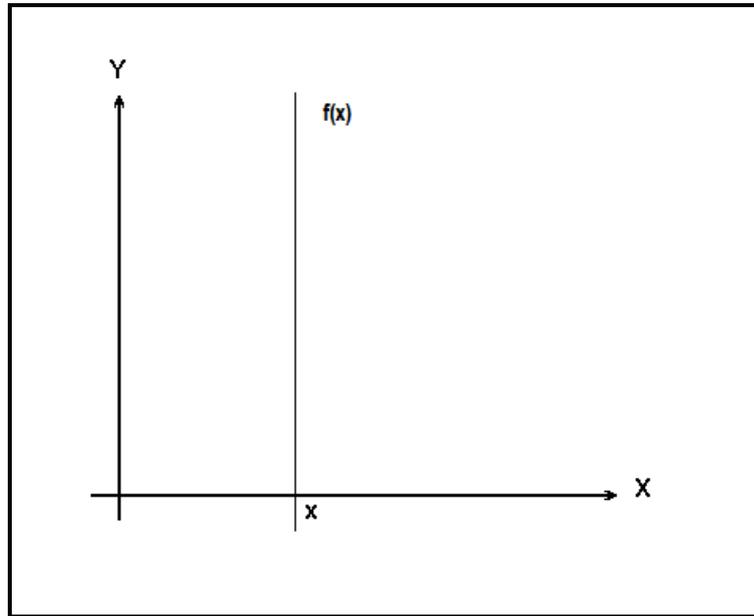
**Fig. N° 8 -Función de distribución marginal de X**



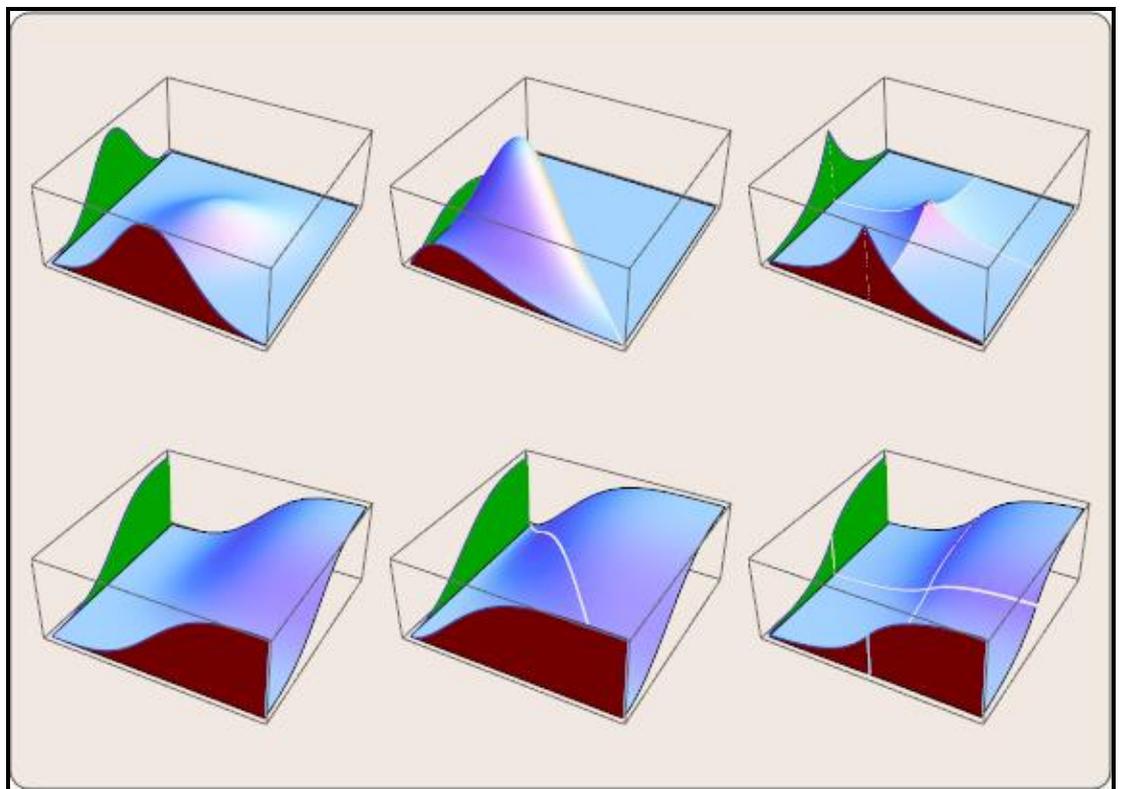
**Fig. N° 9 -Función de distribución marginal de Y**



**Fig. N° 10 -Función de densidad marginal de Y**



**Fig. N° 11 -Función de densidad marginal de X**



**Fig. N° 12 –Funciones marginales**

Si interesa obtener la probabilidad en un intervalo para una de las variables puede obtenerse a partir de la función conjunta o a partir de las marginales:

$$P(a < X < b) = \int_a^b f(x) dx \quad (20)$$

$$P(a < X < b) = \int_a^b \int_{-\infty}^{\infty} f(x,y) dy dx$$

## Funciones de probabilidad condicionales

Queremos conocer cómo se distribuye una de las variables cuando se imponen condiciones a la otra variable. Recordando el concepto de sucesos dependientes:

$$P(A/B) = \frac{P(AB)}{P(B)} \quad P(B/A) = \frac{P(AB)}{P(A)}, \quad \text{con } P(A) \text{ y } P(B) \neq 0$$

En el caso de variable continua la probabilidad de que alguna de las variables tome un valor específico, no tiene sentido:

$$P(X = x) = 0 \quad P(Y = y) = 0$$

Función de densidad condicionada de la v.a. continua  $X$  dado  $Y=y$ :

$$f(X/Y=y) \begin{cases} \frac{f(x,y)}{f(y)} & f(y) > 0 \\ 0 & \text{en el resto} \end{cases} \quad (21)$$

Función de densidad condicionada de la v.a. continua  $Y$  dado  $X=x$ :

$$f(Y/X=x) \begin{cases} \frac{f(x,y)}{f(x)} & f(x) > 0 \\ 0 & \text{en el resto} \end{cases} \quad (22)$$

Puede demostrarse que las anteriores son funciones de densidad, por ejemplo:

$$\int_{-\infty}^{\infty} f(X/Y)dx = \int_{-\infty}^{\infty} \frac{f(x,y)}{f(y)} dx = \frac{1}{f(y)} \int_{-\infty}^{\infty} f(x,y)dx = \frac{f(y)}{f(y)} = 1$$

Función de distribución condicionada de la v.a. continua X dado Y=y<sub>0</sub>:

$$F(x/y) = \int_{-\infty}^x f(x/y)dx = \frac{\int_{-\infty}^x f(x,y)dx}{f(y)} \quad (23)$$

Función de distribución condicionada de la v.a. continua Y dado X=x<sub>0</sub>:

$$F(y/x) = \int_{-\infty}^y f(y/x)dy = \frac{\int_{-\infty}^y f(x,y)dy}{f(x)} \quad (24)$$

## INDEPENDENCIA DE VARIABLES ALEATORIAS

Cuando se tienen dos variables aleatorias distribuidas conjuntamente, puede interesar saber si las variables son dependientes o independientes, ya que puede tener importancia para determinar si es correcto predecir el valor de una basándose en el valor de la otra. Recordar también aquí el concepto de independencia de sucesos:

$$P(AB) = P(A).P(B)$$

$$P(X \leq x; Y \leq y) = P(X \leq x).P(Y \leq y) \quad (25)$$

$$F(x,y) = F(x).F(y)$$

Esto solo es posible de expresar si las variables son independientes. Las variables aleatorias son independientes *si y sólo si* la función conjunta puede expresarse como el producto de sus marginales. Si la variable bidimensional es discreta se expresa de la siguiente forma:

$$P(X = x_i; Y = y_j) = P(X = x_i).P(Y = y_j)$$

$$f(x_i, y_j) = f(x_i).f(y_j)$$

Por lo tanto puede concluirse que: es *necesario* para que dos variables aleatorias distribuidas conjuntamente sean independientes, que su función conjunta (de densidad o cuantía o de distribución) pueda descomponerse en el producto de sus marginales.

Si la función conjunta está expresada como el producto de sus marginales, es *suficiente* para asegurar que las variables son independientes.

Esta condición es válida tanto para el caso continuo como discreto.

## *Ejemplos de variable aleatoria bidimensional*

### *Variable continua*

Considérense dos caudales X e Y en dos corrientes distintas pero en el mismo día. El interés en estudiar el comportamiento conjunto puede originarse debido a que ambas corrientes alimentan al mismo reservorio. La función conjunta está dada por:

$$f(x,y) = \begin{cases} c \cdot \frac{4000-x}{4000} & 0 \leq x \leq 4000 \quad 0 \leq y \leq 2000 \\ 0 & \text{en otra parte} \end{cases}$$

La constante c deberá evaluarse haciendo uso de una de las propiedades de la función de densidad conjunta:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx dy = 1$$

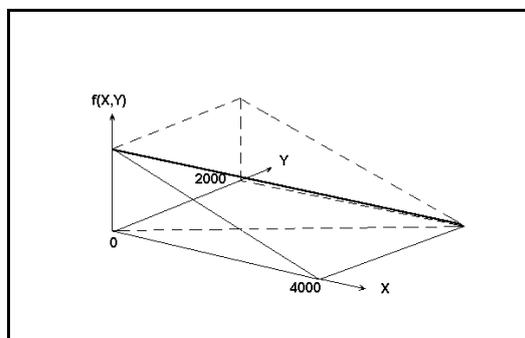
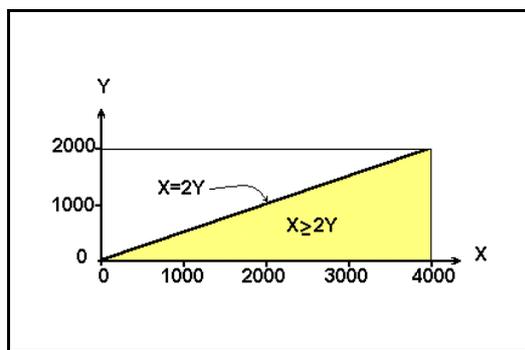
de esta forma:

$$c \cdot \int_0^{4000} \int_0^{2000} \frac{4000-x}{4000} dy dx = 1$$

$$c = 2.5 \cdot 10^{-7}$$

La probabilidad de eventos tales como: el caudal de x es mayor que 2 veces el caudal de y puede interesar al ingeniero que analiza la situación. La probabilidad de este evento es el volumen bajo la función conjunta sobre la región de interés:

$$P(X \geq 2Y) = 2.5 * 10^{-7} \int_0^{4000} \int_0^{0.5x} \frac{1}{4000} \cdot (4000 - x) dy dx$$



Las funciones marginales de X y de Y son:

$$f(x) = 2.5 * 10^{-7} \int_0^{2000} \frac{1}{4000} \cdot (4000 - x) dy = \frac{2.5 * 10^{-7}}{2} \cdot (4000 - x), \quad 0 \leq x \leq 4000$$

$$f(y) = 2.5 * 10^{-7} \int_0^{4000} \frac{1}{4000} \cdot (4000 - x) dx = 2000 * 2.5 * 10^{-7}, \quad 0 \leq y \leq 2000$$

El ingeniero puede utilizar información de Y, el caudal en la corriente menor, para predecir X, caudal en la corriente mayor, con mayor seguridad. Esto puede hacerlo contando con la función condicional  $f(X/Y)$  y la distribución marginal.

$$f(X/Y) = \frac{f(X,Y)}{f(Y)} = \frac{\frac{2.5 * 10^{-7} \cdot (4000 - x)}{4000}}{2.5 * 10^{-7} \cdot 2000}$$

$$f(X/Y) = f(X)$$

Esto dice que el conocimiento de Y no altera la función de X, es decir, no brinda información adicional a la necesidad de predicción.



Universidad Nacional del Litoral  
**Facultad de Ingeniería y Ciencias Hídricas**

## ESTADÍSTICA

**Ingenierías: Recursos Hídricos-Ambiental-Agrimensura**

---

TEORÍA

**Mg. Ing. Susana Vanlesberg**  
Profesor Titular

# UNIDAD 3

## Características de variables aleatorias uni y bidimensionales

## **VARIABLE ALEATORIA UNIDIMENSIONAL**

Como se ha expresado, no es posible predecir con total seguridad el valor futuro de una variable aleatoria, pero sí es posible realizar una descripción completa de su comportamiento a través de sus leyes de probabilidad: función de cuantía, función de densidad, función de distribución.

También es posible condensar en algunos números las características salientes de la distribución de probabilidades de las variables aleatorias. Son fáciles de manejar y también es sencillo obtener estimaciones de ellos.

Estos números son una forma conveniente de cuantificar la ubicación y la forma de una distribución de probabilidades. Se los puede clasificar de la siguiente manera:

- a - Medidas de la tendencia central
- b - Medidas de variabilidad
- c - Medidas de asimetría
- d - Medidas de empinamiento o curtosis

## **MEDIDAS DE LA TENDENCIA CENTRAL**

Se distinguen, dentro de esta clase, las que son *promedios* y las que son *medidas de ubicación*.

### **Promedios**

- a- Esperanza matemática
- b- Media geométrica
- c- Media armónica

### **Medidas de ubicación**

- a- Mediana
- b- Modo
- c- Cuantiles

## PROMEDIOS

### a- Valor esperado o Esperanza Matemática

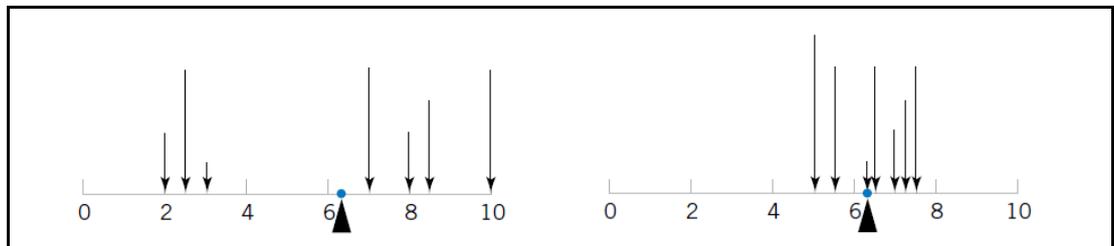
Es uno de los conceptos más importantes y es muy utilizado en la teoría de decisiones, en el análisis de sistemas y en muchos otros campos.

Cálculo para variable discreta y continua:

$$E(X) = \sum_{i=1}^n x_i f(x_i) \quad (1)$$

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx \quad (2)$$

En ella se condensa la información que hay en la función de probabilidad en un solo número. Geométricamente define el centro de gravedad de la distribución de probabilidades de la variable aleatoria.



### Propiedades

1- La esperanza de una constante es igual a la misma constante:

$$E(C) = C$$

$$E(C) = \int_{-\infty}^{\infty} C \cdot f(x) dx = C \cdot \int_{-\infty}^{\infty} f(x) dx = C$$

2- La esperanza de una constante por una función es igual a la constante por la esperanza de la función:

$$E(C X) = C E(X)$$

$$E(CX) = \int_{-\infty}^{\infty} C X f(X) dX = C \int_{-\infty}^{\infty} X f(X) dX = C E(X)$$

3- Esperanza de una función de la variable X:

$$E(a + bX) = a + b E(X)$$

4- Esperanza de una suma de funciones de X:

$$E(g_1(x) \pm g_2(x)) = E(g_1(x)) \pm E(g_2(x))$$

5- Esperanza de un producto de variables aleatorias independientes:

$$E(X.Y) = E(X) . E(Y)$$

### **b - Media geométrica**

Se define como:

$$M_g = \left( \prod_{i=1}^n x_i \right)^{\frac{1}{n}} \quad (3)$$

El logaritmo de la media geométrica es igual al valor esperado de los logaritmos de los  $x_i$ :

$$\log M_g = \sum_{\forall i} f(x_i) \log x_i \quad \text{para variable discreta} \quad (4)$$

$$\log M_g = \int_{-\infty}^{\infty} \log x f(x) dx \quad \text{para variable continua}$$

### **c- Media armónica**

Se define de la siguiente forma:

$$M_H = \frac{1}{\sum_{\forall i} \frac{1}{x_i} f(x_i)} \quad \text{para variable discreta} \quad (5)$$

$$M_H = \frac{1}{\int_{-\infty}^{\infty} \frac{1}{x} f(x) dx} \quad \text{para variable continua}$$

## MEDIDAS DE UBICACIÓN

### a- Mediana

Es una medida de posición de la función de probabilidades. Es el valor de la variable que cumple con la siguiente condición:

$$\int_{-\infty}^{\text{Mediana}} f(x)dx = 0.5$$

$$P(X \leq \text{Mediana}) = 0.5$$

si  $X$  es variable continua

$$\text{o bien } \int_{\text{Mediana}}^{\infty} f(x)dx = 0.5$$

$$P(X > \text{Mediana}) = 0.5$$

$$\text{Prob}(X \leq \text{Mediana}) = \sum_{i=1}^{\text{Mediana}} f(x_i) = 0.5$$

si  $X$  es variable discreta

$$\text{o bien } \text{Prob}(X > \text{Mediana}) = \sum_{\text{Mediana}}^{\infty} f(x_i) = 0.5$$

(6)

### b - Modo

El modo es el valor de la variable que ocurre más frecuentemente. De esta manera, si la variable aleatoria es continua será el valor de  $X$  que maximice a la función de densidad:

$$\frac{df(x)}{dx} = 0 \quad \text{y} \quad \frac{d^2 f(x)}{dx} < 0 \quad (7)$$

y, si la variable es discreta, será el valor de la variable asociado a la máxima probabilidad:

$$\text{Máx } f(x_i) \quad (8)$$

### c - Cuantiles

Se denomina *cuantil de orden p* (siendo p un número perteneciente al intervalo [0,1]) al valor de la variable  $X_p$  que cumple con la siguiente condición:

$$P(X \leq x_p) = p \text{ o bien } P(X > x_p) = 1 - p \quad (9)$$

**Cuartiles.** Dividen a la distribución de probabilidades en cuatro partes, existen tres cuartiles, su expresión es la siguiente:

$$P(X \leq x_{p_i}) = \frac{i}{4}, \text{ para } i = 1, 2, 3 \quad (10)$$

El cuartil de orden 2 corresponde a la mediana.

**Deciles.** Dividen la distribución de probabilidades en 10 partes; por lo tanto, hay 9 deciles:

$$P(X \leq x_{p_i}) = \frac{i}{10}, \text{ para } i = 1, 2, \dots, 9 \quad (11)$$

**Porcentiles.** Dividen la distribución de probabilidades en 100 partes; por lo tanto, hay 99 porcentiles. Son de interés cuando se desea analizar detalladamente la distribución de probabilidades:

$$P(X \leq x_{p_i}) = \frac{i}{100}, \text{ para } i = 1, 2, \dots, 99 \quad (12)$$

## MEDIDAS DE VARIABILIDAD

Estas medidas acompañan a las medidas de tendencia central e indican la dispersión o aleatoriedad en el comportamiento de la variable aleatoria.

a - Rango

b - Varianza - Desvío

c - Coeficiente de variabilidad

### a - Rango

Es simplemente la diferencia entre el mayor y el menor valor de la variable; es una cantidad que no aporta mucha información, ya que solamente se considera un par de números. Suele variar desde  $-\infty$  a  $+\infty$  ó de 0 a  $+\infty$ , o bien entre dos números.

### b - Varianza

Es la medida de dispersión más usada.

$$\text{Var}[X] = \sigma_x^2 = E[X - E[X]]^2$$

$$\text{Var}[X] = \int_{-\infty}^{+\infty} (x - E[X])^2 f(x) dx, \text{ si } X \text{ es continua} \quad (13)$$

$$\text{Var}[X] = \sum_{\forall i} (x_i - E[X])^2 f(x_i), \text{ si } X \text{ es discreta}$$

Observar que su expresión responde a la necesidad de una medida que exprese los apartamientos de los valores de la variable respecto a la media que es el valor representativo de la tendencia central.

Esta expresión puede simplificarse y ser expresada en función de los momentos con respecto al origen:

$$\begin{aligned} E[X - E[X]]^2 &= E[X^2 - 2XE[X] + E^2[X]] = \\ &= E[X^2] - 2E[X]E[X] + E[E^2[X]] = \\ &= E[X^2] - 2E^2[X] + E^2[X] = \\ &= E[X^2] - E^2[X] = \alpha_2 - \alpha_1^2 \end{aligned} \quad (14)$$

**Momentos:** se los define como los promedios de distintas potencias de la variable aleatoria (tanto discreta como continua):

$$\alpha_k = E[X^k] = \sum_{\forall i} x_i^k \cdot f(x_i) \quad (15)$$

$$\alpha_k = E[X^k] = \int_{-\infty}^{+\infty} x^k f(x) dx$$

Estas expresiones corresponden a los momentos k-ésimos del área de la función de probabilidades *con respecto al origen*.

Es posible definir también los momentos de áreas con respecto a otro punto que no sea el origen. En particular, momentos con respecto a la media se denominan *momentos centrados*.

$$\mu_k = E[(X - E[X])^k] = \int_{-\infty}^{+\infty} (x - E[X])^k f(x) dx, \text{ si } X \text{ es continua} \quad (16)$$

$$\mu_k = E[(X - E[X])^k] = \sum_{\forall i} (x_i - E[X])^k \cdot f(x_i), \text{ si } X \text{ es discreta}$$

El momento de orden 1 es igual a 0:

$$\mu_1 = E[x - E[X]] = E[X] - E[E[X]] = E[X] - E[X] = 0$$

El momento centrado de orden 2 define a la **varianza de X**.

### Propiedades

1- La varianza de una constante es igual a cero:

$$\text{Var}(C) = 0$$

$$E[C - E(C)]^2 = E[C - C]^2 = E(0) = 0$$

2- La varianza de una constante por X es igual a la constante al cuadrado por la varianza de la variable X:

$$\text{Var}(CX) = C^2 \text{Var}(X)$$

$$\begin{aligned} E[CX - E(CX)]^2 &= E[CX - CE(X)]^2 = \\ &= E\{[C \cdot (X - E(X))]^2\} = E\{C^2 \cdot [X - E(X)]^2\} = \\ &= C^2 E[X - E(X)]^2 = C^2 \text{Var}(X) \end{aligned}$$

3- Varianza de una función de X:

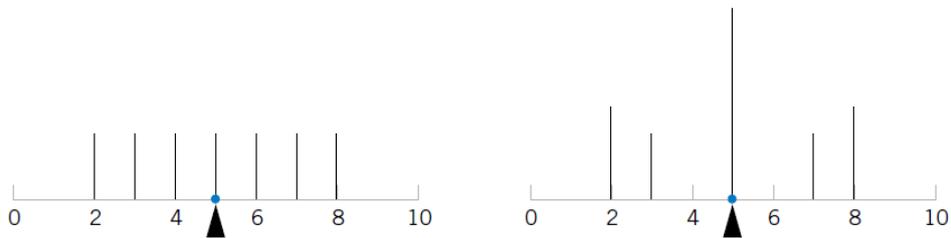
$$\text{Var}(a + bX) = \text{Var}(a) + \text{Var}(bx) = 0 + b^2 \text{Var}(X)$$

### Desvío Estándar

Para algunas aplicaciones suele ser más conveniente utilizar una medida de variabilidad en las mismas unidades de la variable; para eso se define el *desvío estándar* como la raíz cuadrada positiva de la varianza:

$$\sigma_x = +\sqrt{\text{Var}(X)} \tag{17}$$

$$\sigma_x = +\sqrt{\alpha_2 - \alpha_1^2}$$



**Las distribuciones de la Figura anterior son distintas, pero observar que a pesar de eso poseen la misma media y la misma varianza. Esto indica que se debe tener cuidado en su interpretación.**

### Coefficiente de Variabilidad

La varianza es una medida de dispersión que está influenciada por la magnitud de los valores de dicha variable y por la media. Para evitar esta influencia una medida relativa de dispersión que expresa la dispersión de la variable respecto del tamaño de la variable, midiendo el tamaño de la variable por su valor esperado, es el coeficiente de variación:

$$Cv = \frac{\sigma}{\mu} (\%) \quad (18)$$

Este coeficiente expresa la dispersión de una v.a. respecto a su media y es muy útil para comparar dos distribuciones de probabilidad.

El coeficiente de variación no tendrá sentido cuando la variable X tome valores positivos y negativos (pues puede ocurrir que la media quede compensada por los valores positivos y negativos y no refleje el tamaño de X). Con lo que el coeficiente de variación sólo tendrá sentido cuando X sea una variable que tome sólo valores positivos.

Otra medida de dispersión asociada a los cuartiles es el Recorrido intercuartílico:

$$RQ = Q3 - Q1 \quad (19)$$

Dentro de este intervalo intercuartílico se encuentran el 50% de los valores centrales de la variable.

### **Teorema de Tchebycheff**

La desigualdad de Tchebycheff ofrece ayuda en este punto ya que brinda límites para las probabilidades de apartamiento de la variable aleatoria un determinado número de veces el desvío respecto a su valor medio. Es de ayuda en la interpretación del desvío estándar.

Para describir completamente el comportamiento de una variable aleatoria es necesario conocer la función de probabilidad de la misma. Pero si se conocen la media y la desviación estándar es posible obtener algún tipo de información acerca de la distribución de probabilidades y obtener cotas para calcular probabilidades de apartamiento respecto a la media.

**Enunciado:** una variable seleccionada al azar de una distribución de probabilidades (que se desconoce) se desviará más que  $h \sigma$  de su media  $\mu$  con una probabilidad menor o igual que  $1/h^2$  siendo  $h > 0$ .

$$P(|X - \mu| \geq h \sigma) \leq \frac{1}{h^2} \quad (20)$$

o también en forma complementaria:

$$P(|X - \mu| < h\sigma) \geq 1 - \frac{1}{h^2} \quad (21)$$

Para demostrar este teorema se partirá de la expresión del momento centrado de orden 2: la varianza, y como se quiere hacer notar el papel importante de la varianza como medida de dispersión, se considerará su expresión para una variable continua como caso particular:

$$\sigma^2 = \mu_2 = E[X - E(X)]^2 = \int_{-\infty}^{\infty} [X - E(X)]^2 f(x) dx$$

La expresión planteada por (18) corresponde a aquellos valores que se apartan más de  $h$  veces  $\sigma$  del valor medio  $\mu$ .



La probabilidad planteada puede ahora expresarse de la siguiente manera:

$$P(\mu - h\sigma \geq X \geq \mu + h\sigma) \leq \frac{1}{h^2}$$

$$1 - \int_{\mu - h\sigma}^{\mu + h\sigma} f(x) dx \leq \frac{1}{h^2}$$

Considerando a  $\sigma^2$  entre los mismos límites:

$$\sigma^2 = \int_{-\infty}^{\mu - h\sigma} (X - \mu)^2 f(x) dx + \int_{\mu - h\sigma}^{\mu + h\sigma} (X - \mu)^2 f(x) dx + \int_{\mu + h\sigma}^{\infty} (X - \mu)^2 f(x) dx$$

(I) (II)

Para tratar de encontrar relación entre ambas expresiones, se considerarán aquellos valores que se aparten de  $\mu$  más de  $h\sigma$ , que son los que caen en la zona delimitada por  $(-\infty, \mu-h)$ ,  $(\mu+h\sigma, \infty)$  Reemplazando  $(x-\mu)^2$  por  $h^2\sigma^2$  y trabajando con las integrales que corresponden (I y II) se obtiene lo siguiente:

$$\int_{-\infty}^{\mu-h\sigma} h^2\sigma^2 f(x)dx + \int_{\mu+h\sigma}^{\infty} h^2\sigma^2 f(x)dx \leq \sigma^2$$

$$h^2\sigma^2 \left[ \int_{-\infty}^{\mu-h\sigma} f(x)dx + \int_{\mu+h\sigma}^{\infty} f(x)dx \right] \leq \sigma^2$$

$$P(X \leq \mu - h\sigma) + P(X \geq \mu + h\sigma) \leq \frac{\sigma^2}{h^2\sigma^2}$$

$$P(|X - \mu| \geq h\sigma) \leq \frac{1}{h^2}$$

El significado importante de usar la desigualdad de Tchebycheff es que puede ser utilizada para cualquier variable cuando no se tiene información sobre su función, *indica además, cómo puede ser usado el desvío como medida de variación.*

## MEDIDAS DE FORMA

### MEDIDAS DE ASIMETRÍA

De igual forma que la media y la varianza miden la ubicación y dispersión de una distribución, los momentos más altos miden otras propiedades de la misma.

Necesitamos definir una serie de medidas que permitan cuantificar, en la medida de lo posible, la forma de la distribución, es decir, tendremos que definir unas medidas que den información sobre la función de probabilidad o sobre la función de densidad, de forma que podamos cuantificar el grado de simetría y el de apuntamiento o de aplastamiento, bien de la función de probabilidad, o bien de la función de densidad.

El tercer momento respecto a la media es usado para determinar si una distribución es simétrica o asimétrica.

Si la *distribución es simétrica*, como las desviaciones están elevadas al cubo, las positivas y negativas tienden a anularse y, por lo tanto,  $\mu_3 = 0$ .

Si la *distribución es asimétrica a la derecha*,  $\mu_3 > 0$ .

Si la *distribución es asimétrica a la izquierda*,  $\mu_3 < 0$ .

$\mu_3$ , tomado como valor aislado, no es una buena medida de la asimetría, ya que tiene las mismas unidades que la variable; por eso es que se define una medida relativa denominada *coeficiente de asimetría*:

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \quad (22)$$

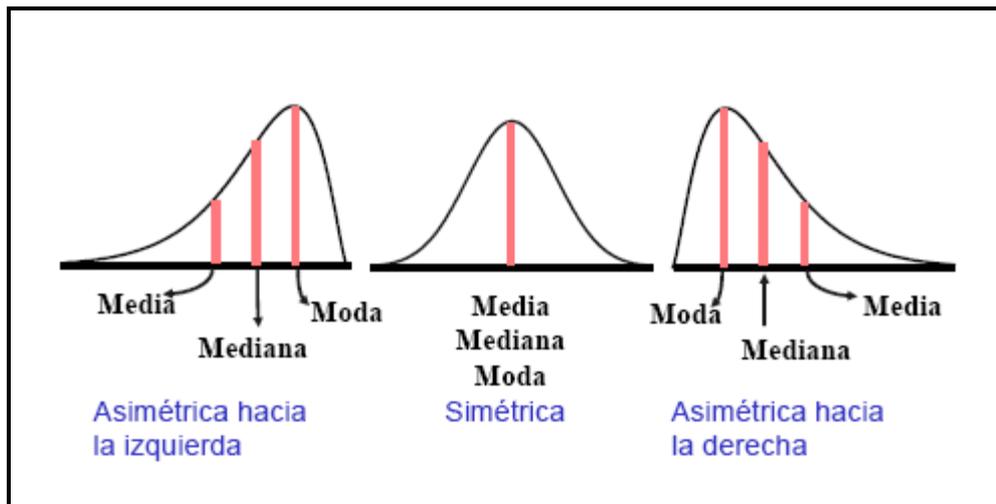
En función de los momentos respecto al origen:

$$\begin{aligned} \mu_3 &= E(X - E[X])^3 = E[X^3 - 3X^2 E[X] + 3X E^2[X] - E^3[X]] = \\ &= E[X^3] - 3E[X^2]E[X] + 3E^3[X] - E^3[X] = \\ &= \alpha_3 - 3\alpha_2\alpha_1 + 3\alpha_1^3 - \alpha_1^3 \\ \mu_3 &= \alpha_3 - 3\alpha_2\alpha_1 + 2\alpha_1^3 \\ \gamma_1 &= \frac{\alpha_3 - 3\alpha_2\alpha_1 + 2\alpha_1^3}{(\alpha_2 - \alpha_1^2)^{3/2}} \quad (23) \end{aligned}$$

Se cumplen también las siguientes relaciones aproximadas:

$$As = \frac{E(X) - \text{Modo}}{\sigma(X)} \quad \text{bien} \quad As = \frac{3[E(X) - \text{Mediana}]}{\sigma(X)} \quad (24)$$

La distribución será simétrica si  $As=0$ ; asimétrica a la derecha si  $As>0$ ; y asimétrica a la izquierda si  $As<0$ .



## MEDIDAS DE CURTOSIS

Una cuarta propiedad de las variables aleatorias se basa en el momento centrado de cuarto orden, que permite evaluar el empinamiento o aplastamiento de la distribución de probabilidades comparada con una curva tomada como modelo (la curva normal).

Una medida relativa para independizarse de las unidades:

$$\gamma_2 = \frac{\mu_4}{\sigma^4} \quad (25)$$

En función de los momentos respecto al origen:

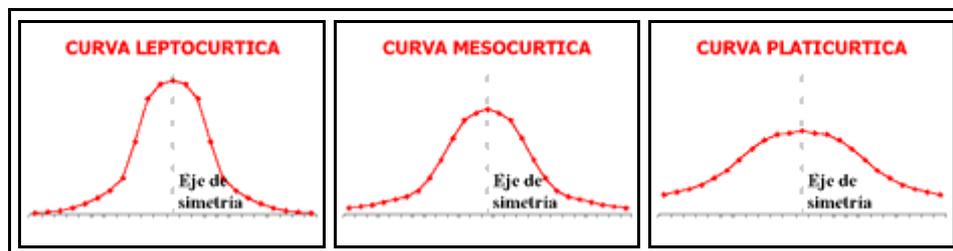
$$\mu_4 = E[X - E[X]]^4 = E[X^4 - 4X^3 E[X] + 6X^2 E^2[X] - 4X E^3[X] + E^4[X]] =$$

$$= E[X^4] - 4E[X^3]E[X] + 6E[X^2]E^2[X] - 4E[X]E^3[X] + E^4[X]$$

$$\mu_4 = \alpha_4 - 4\alpha_3\alpha_1 + 6\alpha_2\alpha_1^2 - 3\alpha_1^4 \quad (26)$$

$$\gamma_2 = \frac{\alpha_4 - \alpha_3\alpha_1 + 6\alpha_2\alpha_1^2 - 3\alpha_1^4}{(\alpha_2 - \alpha_1^2)^2} \quad (27)$$

La Curtosis de la curva Normal tomada como referencia es igual a 3, y se la denomina distribución *mesocúrtica*. Para distribuciones que presenten mayor concentración de probabilidad cerca de la media, mayor que en la Normal, la curtosis será mayor que 3 y se denominará *leptocúrtica*. En caso que la concentración alrededor de la media sea menor que en la Normal la curtosis será menor que 3 y la distribución se dice *platicúrtica*.



## CARACTERÍSTICAS DE VARIABLES ALEATORIAS DISTRIBUIDAS CONJUNTAMENTE

Cuando dos o más variables aleatorias deben ser tratadas simultáneamente, su comportamiento se describe a través de su distribución de probabilidades conjuntas. Los conceptos anteriores sobre características se extienden sencillamente a este caso.

### Momentos

Se parte de considerar una función de las variables X e Y de la siguiente forma:

$$g(X,Y) = X^l Y^n \quad (28)$$

Para obtener el momento conjunto de orden l,n se obtiene la esperanza de la expresión anterior, desarrollada para el caso de una variable bidimensional continua:

$$E(X^l, Y^n) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^l y^n f(x, y) dx dy \quad (29)$$

Los momentos más importantes son los de orden (1,0), (0,1), (2,0) y (0,2).

En el caso que  $g(X,Y) = X$ , se obtiene la esperanza de X:

$$E[g(X,Y)] = E(X) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy \quad (30)$$

$$E(X) = \alpha_{1,0} = \int_{-\infty}^{\infty} x dx \int_{-\infty}^{\infty} f(x, y) dy$$

Como sucede en el caso unidimensional, en ambos casos los resultados son números y para que exista el valor esperado es necesario que la correspondiente serie o integral de la definición sean absolutamente convergentes, es decir, será necesario que  $E[g(X,Y)]$  sea finito.

Para obtener la esperanza de Y se procede de forma análoga:

$$E(Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy = \int_{-\infty}^{\infty} y dy \int_{-\infty}^{\infty} f(x, y) dx$$

$$E(Y) = \alpha_{0,1} = \int_{-\infty}^{\infty} y f(y) dy \quad (31)$$

$\alpha_{1,0}$  y  $\alpha_{0,1}$  establecen el centro de masa de la distribución de probabilidades.

Por analogía, los momentos de segundo orden  $\alpha_{2,0}$  y  $\alpha_{0,2}$  corresponden al momento de inercia respecto a cada uno de los ejes: x, y .  
 $\alpha_{1,1}$  corresponde al momento de inercia producto con respecto a ambos ejes.

### Momentos centrados

Como en el caso de una variable, y como en mecánica, el momento centrado más usado es el de segundo orden. Tomando ahora una función de ambas variables igual a:

$$g(X,Y) = [X - E(X)]^l \cdot [Y - E(Y)]^n \quad (32)$$

hallando la esperanza de esta función, se obtiene :

$$E[g(x,y)] = E \{ [x - E(X)]^l \cdot [y - E(Y)]^n \} \quad (33)$$

Esta expresión se denomina momento centrado de orden l,n de las variables aleatorias x, y.

De la misma manera que para el caso unidimensional el momento centrado de orden 1,0 o 0,1 es igual a cero.

Los momentos más usados son los de orden 2,0, 0,2 y 1,1. Los momentos de orden 2,0 y 0,2 llevan a obtener las varianzas marginales:

$$E[g(x,y)] = E \{ [x - E(X)]^2 \cdot [y - E(Y)]^0 \} \quad (34)$$

Puede ser también expresada en función de los momentos respecto al origen, desarrollando la expresión anterior:

$$\sigma^2(X) = \alpha_{2,0} - \alpha_{1,0}^2 \quad (35)$$

Un resultado similar se obtiene para la varianza de Y:

$$\sigma^2(Y) = \alpha_{0,2} - \alpha_{0,1}^2 \quad (36)$$

### Covarianza

Un momento muy importante se obtiene de considerar l=1 y n=1. Este momento se denomina *covarianza*.

La covarianza permite dar una medida de la dependencia lineal entre X e Y.

Covarianza positiva: Cuando los valores de X crecen los valores de Y también tiende a crecer.

Covarianza negativa: Cuando los valores de X crecen los valores de Y tiende a decrecer.

La covarianza no es una medida de las relaciones o dependencias entre las dos variables X e Y, sino que únicamente es una medida de la fuerza de la relación lineal entre X e Y.

$$\begin{aligned} \text{Cov}(X,Y) &= \sigma_{x,y} = E\{ [X - E(X)] [y - E(Y)] \} = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x - E(X)] [y - E(Y)] f(x, y) dx dy \end{aligned} \quad (37)$$

Puede ser expresada en función de los momentos respecto al origen:

$$\sigma_{x,y} = \alpha_{1,1} - \alpha_{1,0} \cdot \alpha_{0,1} \quad (38)$$

Un inconveniente para utilizar la covarianza, es que la magnitud la covarianza no es significativa, pues depende de las unidades de medida utilizadas en X e Y.

### Coefficiente de correlación

Una versión normalizada de la covarianza es el denominado *coeficiente de correlación*  $\rho(x,y)$ , y se obtiene dividiendo la covarianza por el producto de sus desviaciones:

$$\begin{aligned} \rho_{x,y} &= \frac{\text{cov}_{x,y}}{\sigma_x \cdot \sigma_y} = \frac{\mu_{1,1}}{\sqrt{\mu_{2,0} \cdot \mu_{0,2}}} \\ &-1 \leq \rho \leq 1 \end{aligned} \quad (39)$$

Este coeficiente merece algunos comentarios;

-de la misma manera que la media y la varianza, el coeficiente de correlación es útil cuando la ley de probabilidades no se conoce en forma completa; en este caso pares de observaciones de ambas variables pueden ayudar a estimar el coeficiente de correlación y poder obtener alguna conclusión acerca de su comportamiento a través de su valor.

-de la definición de covarianza se ve que valores positivos de ella resultarán de pares de valores altos de x con valores altos de y o valores pequeños de x y pequeños de y; mientras que valores negativos de la covarianza se darán con la asociación de valores pequeños de una con valores grandes de la otra. En ambos casos es posible decir que existe al menos alguna dependencia estocástica entre ambas.

-si dos variables aleatorias son independientes, su covarianza y por lo tanto su coeficiente de correlación serán iguales a cero. La independencia implica que  $f(x,y)$  puede descomponerse en el producto de las marginales  $f(x).f(y)$  y entonces:

$$\begin{aligned} & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x - E(X)].[y - E(Y)]. f(x). f(y) dx dy = \\ & = \int_{-\infty}^{\infty} [x - E(X)]. f(x) dx \int_{-\infty}^{\infty} [y - E(Y)]. f(y) dy = \\ & = E[x - E(X)].E[y - E(Y)] = \mu_{1,0} \cdot \mu_{0,1} \end{aligned}$$

$$\text{siendo } \mu_{1,0} = 0 \text{ y } \mu_{0,1} = 0$$

-lo inverso no es cierto, o sea si el coeficiente de correlación o la covarianza son cero esto no implica que las variables sean independientes, puede haber alta dependencia estadística y ser  $\rho=0$ . El coeficiente de correlación es una medida de dependencia lineal entre dos variables aleatorias.

### **Esperanza de una suma o de una diferencia de variables aleatorias**

Si interesa encontrar el valor esperado de la suma o diferencia de dos variables aleatorias X e Y , cada una con esperanza conocida E(X) y E(Y), es directamente la suma o diferencia de sus esperanzas:

$$E(X \pm Y) = E(X) \pm E(Y) \quad (40)$$

Esto se demuestra fácilmente partiendo del concepto de esperanza:

$$E(X \pm Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x \pm y) f(x, y) dx dy =$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dy dx \pm \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy =$$

$$\int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f(x, y) dy dx \pm \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f(x, y) dx dy =$$

$$\int_{-\infty}^{\infty} x f(x) dx \pm \int_{-\infty}^{\infty} y f(y) dy =$$

$$E(X) \pm E(Y)$$

Este resultado puede extenderse al caso de más de dos variables y es válido tanto para variables aleatorias independientes como dependientes. La demostración fue hecha para el caso de variables continuas pero es similar para el caso discreto.

### Esperanza del producto de variables aleatorias

Si se considera ahora una función de dos variables aleatorias X e Y :

$$Z = X.Y$$

Entonces

$$E(Z) = E(XY) = Cov(XY) + E(X).E(Y)$$

$$Cov(XY) = E\{ [X - E(X)].[Y - E(Y)] \} =$$

$$= E\{ XY - X.E(Y) - E(X).Y + E(X).E(Y) \} =$$

$$= E(XY) - E(X).E(Y) - E(X).E(Y) + E(X).E(Y)$$

Si las variables aleatorias son independientes, la covarianza es igual a cero; luego

$$E(XY) = E(X).E(Y)$$

(41)

$$E(XY) = Cov(XY) + E(X).E(Y)$$

### Varianza de una suma y de una diferencia de variables aleatorias

$$Var(X + Y) = E\{[(X - E(X)) + (Y - E(Y))]^2\}$$

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \cdot \text{Cov}(XY) \quad (42)$$

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - 2 \cdot \text{Cov}(XY) \quad (43)$$



Universidad Nacional del Litoral  
Facultad de Ingeniería y Ciencias Hídricas

## ESTADÍSTICA

**Ingenierías:** Recursos Hídricos-Ambiental-Agrimensura

---

### TEORÍA

*Mg.Ing. Susana Vanlesberg*  
Profesor Titular

# UNIDAD 4

## Modelos Probabilísticos

### Variable Discreta

La aplicación de la teoría de probabilidad en situaciones reales, concretas, que presentan una regularidad en el desarrollo, ha originado una serie de modelos que permiten resolver esas situaciones. Los ingenieros realizan supuestos respecto del problema a resolver, esto los lleva a descripciones análogas y a formas matemáticamente iguales a las de los modelos probabilísticos.

En este capítulo no sólo se presentan los modelos más usados en la práctica, sino que también se brindan algunas nociones de los mecanismos por los cuales se ha originado cada distribución. Esto último es de suma importancia para un ingeniero, ya que la existencia de tales mecanismos pueden describir una situación física de su interés, y ésta es una razón más importante que la buena aproximación matemática de algún modelo.

## MODELOS DE VARIABLE ALEATORIA DISCRETA

Cuando en el desarrollo de los modelos se mencionen pruebas, éstas se referirán a observaciones sistemáticas de algún fenómeno natural.

### Modelo Bernoulli

Tal vez la situación más común que se presenta es aquella en la que los resultados de los experimentos pueden separarse en dos categorías mutuamente excluyentes: *éxito* o *fracaso*; por ejemplo llueve o no, mido o no mido, está contaminado o no, etc.

Puede entonces definirse la variable aleatoria  $x$  de tipo Bernoulli y asignársele valores (arbitrarios pero muy prácticos) a los eventos antes mencionados:

$x=0$       fracaso

$x=1$       éxito

La función de probabilidad de  $x$  es simplemente:

$$f(x) = \begin{cases} p, & x = 1 \\ 1 - p, & x = 0 \end{cases} \quad (1)$$

$p$  probabilidad de éxito.

Esta función deberá cumplir con las condiciones de una función de probabilidad:

$$\sum_{x_i=0}^{x_i=1} f(x_i) = 1$$

$$p + (1 - p) = p + 1 - p = 1$$

Las características de este modelo:

$$E(x) = \sum_{i=0}^1 x_i f(x_i) = 0(1 - p) + 1p = p \quad (2)$$

$$\begin{aligned} V(x) &= \sum_{-x_i} (x_i - E(x))^2 f(x_i) = \\ &= (0 - p)^2(1 - p) + (1 - p)^2 \cdot p = \\ &= p^2(1 - p) + (1 - p)^2 \cdot p = \\ &= p(1 - p)(p + 1 - p) \end{aligned}$$

$$V(x) = p(1 - p) \quad (3)$$

## Modelo Binomial

Se realizan una serie de pruebas de tipo Bernoulli cuyos resultados son mutuamente independientes y si la probabilidad de éxito permanece invariable en todas ellas se origina un nuevo modelo denominado **BINOMIAL**.

Para determinar la distribución correspondiente se analiza el número total de éxitos en  $n$  pruebas Bernoulli cada una con probabilidad favorable igual a  $p$ .

Se consideran por ejemplo 3 pruebas y se analizarán las probabilidades de éxito:

- Ningún éxito:  $x=0$

0 ; 0 ; 0 ;

$(1 - p)^3$

- Un éxito:  $x=1$

1 ; 0 ; 0 ;   ó   0 ; 1 ; 0 ;   ó   0 ; 0 ; 1

cada secuencia es un evento, que entre sí son mutuamente excluyentes cuya probabilidad es :

$p(1-p)^2$ ; de esta manera la probabilidad de un éxito es :

$$3p(1-p)^2$$

- Dos éxitos:  $x=2$

1; 1; 0;    ó    1; 0; 1;    ó    0; 1; 1;

nuevamente cada secuencia es un evento que es excluyente de los restantes:

$p^2(1-p)$  y el total :

$$3p^2(1-p)$$

- Tres éxitos:  $x=3$

1; 1; 1;            p; p; p

$$p^3$$

Esto permite generalizar y obtener la función de probabilidad de este modelo:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

(4)

$$\text{coeficiente binomial} \binom{n}{x} = \frac{n!}{x!(n-x)!}$$

El coeficiente binomial tiene en cuenta el número de formas en que exactamente  $x$  éxitos se presenten en  $n$  pruebas. Los parámetros de este modelo son  $n$  y  $p$ ;  $n$  entero y  $p$  un número comprendido entre 0 y 1.

El modelo se denomina **BINOMIAL** porque puede considerarse como el desarrollo del binomio  $(p + q)^n$ .

La función debe cumplir las condiciones de una función de probabilidad:

$$\begin{aligned} \sum_{x_i} f(x_i) &= 1 \\ \sum_{x_i} \binom{n}{x_i} p^{x_i} q^{n-x_i} &= \\ &= \sum_{x_i=0}^n \frac{n!}{x_i!(n-x_i)!} p^{x_i} q^{n-x_i} = \\ &= \frac{n!}{0!n} p^0 q^n + \frac{n!}{1!(n-1)!} p q^{n-1} + \dots + \frac{n!}{n!0!} p^n q^0 = \\ &= q^n + \frac{n(n-1)!}{(n-1)!} p q^{n-1} + \dots + p^n = (p+q)^n = 1^n = 1 \end{aligned}$$

La función de distribución puede obtenerse aplicando la definición:

$$F(x) = P(X \leq x) = \sum_{\forall x_i \leq x} f(x_i) \quad (5)$$

### Características del modelo

$$\begin{aligned} E(X) &= \sum_{x_i=0}^n x_i f(x_i) = \\ &= 0 \cdot \frac{n!}{0!(n-0)!} p^0 q^n + 1 \cdot \frac{n!}{1!(n-1)!} p^1 q^{n-1} + \dots + n \cdot \frac{n!}{n!(n-n)!} p^n q^0 = \\ &= 0 + \frac{n(n-1)!}{(n-1)!} p q^{n-1} + \dots + n p^n q^0 = np(q^{n-1} + \dots + (n-1)p^{n-2}q + p^{n-1}) = \\ &= np(p+q)^{n-1} \end{aligned}$$

$$E(x) = np \quad (6)$$

$$\text{Var}(X) = E[X - E(X)]^2 = E(X^2) - E^2(X)$$

$$E(X^2) = \sum_{x_i=0}^n \binom{n}{x_i} p^{x_i} q^{n-x_i} x_i^2 =$$

ya que el primer termino es cero  
la suma se toma desde 1

$$\begin{aligned} &= np \sum_{x_i=1}^n \frac{(n-1)!}{x_i!(n-x_i)!} p^{x_i-1} q^{n-x_i} x_i^2 \\ &= np \sum_{x_i=1}^n \frac{(n-1)!}{x_i(x_i-1)!(n-x_i)!} p^{x_i-1} q^{n-x_i} x_i^2 \end{aligned}$$

haciendo  $x_i = x_j + 1$ , para  $x_i = 1, x_j = 0, x_i = n, x_j = n-1$   
 $x_i! = (x_i - 1)! x_i = x_j!(x_j + 1)$

$$\begin{aligned} &np \sum_{x_j=0}^{n-1} \frac{(n-1)!}{(x_j)!(x_j+1)(n-x_j-1)!} p^{x_j} q^{n-1-x_j} (x_j+1) = \\ &np \left[ \sum_{x_j=1}^{n-1} \binom{n-1}{x_j} p^{x_j} q^{n-1-x_j} x_j + \sum_{x_j=0}^{n-1} \binom{n-1}{x_j} p^{x_j} q^{n-1-x_j} \right] \end{aligned}$$

El 2do termino de la suma es 1; se analiza el 1º,

$$\text{ya que } \binom{n-1}{x_j} x_j = \frac{(n-1)! x_j}{x_j!(n-1-x_j)!} = \frac{(n-2)!(n-1)x_j}{(x_j-1)! x_j(n-2-(x_j-1))!}$$

sacando factor comun  $(n-1)p$ , se consigue

$$\sum_{x_j=1}^{n-1} \binom{n-2}{x_j-1} p^{x_j-1} q^{n-1-x_j}$$

haciendo  $x_j = x_k + 1$ ,  $x_k$  varía de 0 a  $n-2$  con lo cual

$$\sum_{x_k=0}^{n-2} \binom{n-2}{x_k} p^{x_k} q^{n-2-x_k} = 1$$

$$E(X^2) = np[(n-1)p + 1] = n^2 p^2 - np^2 + np$$

$$\text{Var}(X) = npq \quad (7)$$

Estos mismos resultados podrían haberse obtenido de forma mucho más sencilla por considerar a la variable aleatoria  $X$  como la suma de  $n$  variables independientes idénticamente distribuidas como Bernoulli, cuya esperanza es  $p$  y

$$E(X) = E(x_1 + x_2 + \dots + x_n) = \sum_{i=1}^n E(x_i) = n E(x_i) = n p$$

varianza  $p \cdot q$ :

$$Var(X) = \sum_{i=1}^n Var(x_i) = n p(1 - p) = n p q$$

La distribución Binomial ha sido tabulada, tanto su función de cuantía como la de distribución. Para el caso de la función masa existen valores hasta  $p = 0.5$ . En caso de necesitarse  $p > 0.5$ , se debería plantear el problema cambiando el suceso en estudio por su contrario y tener cuidado en la interpretación de los resultados. También en los software que poseen fórmulas estadísticas y en los específicos se pueden encontrar los valores de esta distribución para cualquier valor de  $p$ .

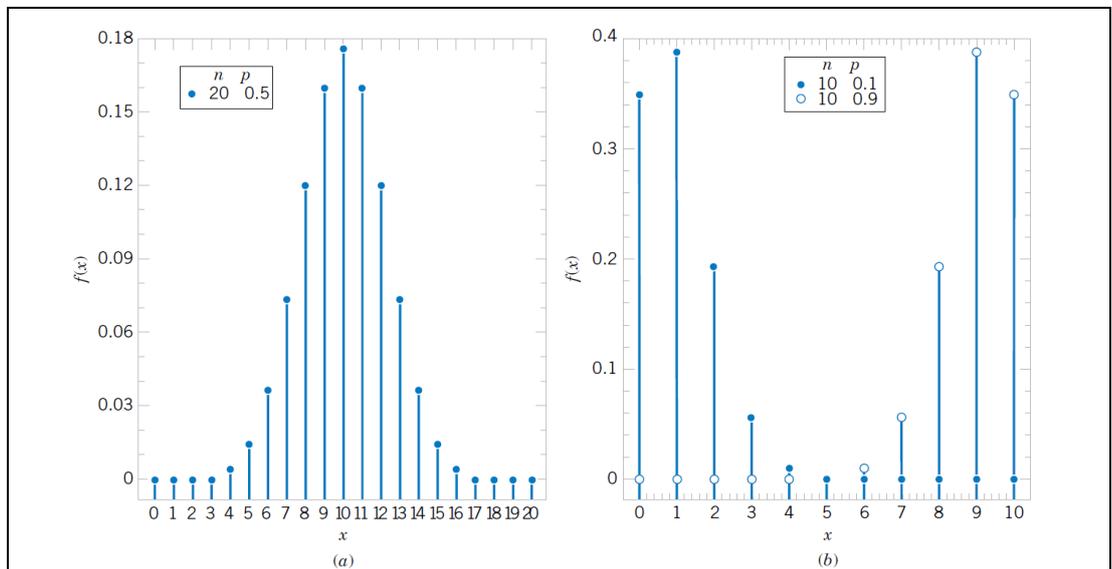


Figura N° 1 - Distribución Binomial para distintos valores de  $n$  y  $p$

### Ejemplo

Cada muestra de agua tiene una probabilidad del 10% de contener un contaminante orgánico particular. Asumir que las muestras son independientes con respecto a la presencia del contaminante. Encontrar la probabilidad de que en las próximas 18 muestras, exactamente 2 contengan el contaminante y el valor medio de muestras contaminadas.

### Solución:

La variable será:

$X =$  " número de muestras que contienen el contaminante"

$$X \sim B(n=18; p=0,1)$$

$$P(X = 2) = \binom{18}{2} 0.1^2 0.9^{16} = 0.284$$

$$E(X) = np = 18 \cdot 0.1 = 1.8 \cong 2$$

## Modelo Geométrico

Continuando con el mismo esquema de pruebas repetidas de tipo Bernoulli, puede interesar conocer en que prueba ocurrirá el primer éxito.

Las pruebas son repetidas, independientes, con dos posibles resultados: ocurrencia o no del evento en estudio, con probabilidad favorable  $p = \text{constante}$ ; se deriva el modelo correspondiente a la variable  $N$ : número de pruebas hasta que ocurre el primer éxito. Este primer éxito se producirá sí y sólo sí en las pruebas anteriores no se produjo, lo cual ocurre con probabilidad igual a  $(1 - p)$ . Luego su función de probabilidad es:

$$P(N = n) = f(n) = (1 - p)^{n-1} p \quad \text{con } n = 1, 2, 3, \dots (8)$$

Este modelo se denomina **GEOMÉTRICO** de parámetro  $p$ .

La función acumulativa se obtiene a partir de la función anterior:

$$F(n) = \sum_{j=1}^n p(j) = \sum_{j=1}^n (1 - p)^{j-1} p =$$

$$= p \left[ \frac{(1 - p)^n - 1}{(1 - p) - 1} \right]$$

$$\left[ \frac{(1 - p)^n - 1}{(1 - p) - 1} \right]$$

*desarrollo de la  
suma de una progresión geométrica*

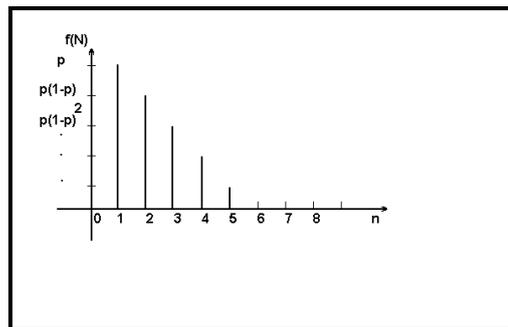
$$F(x) = 1 - (1 - p)^n \quad (9)$$

Esto mismo se puede obtener al considerar la probabilidad de que  $N \leq n$  simplemente como la probabilidad que exista al menos una ocurrencia en  $N$  pruebas:

$$P(N \geq 1) = 1 - P(N < 1) = 1 - P(\text{no haya ocurrencias en } n \text{ pruebas}) =$$

$$F(x) = 1 - (1 - p)^n$$

La gráfica de la función de cuantía:



**Figura N° 2-**Función del modelo Geométrico

### Características

$$E(N) = \sum_{n=1}^{\infty} n p (1-p)^{n-1} = p \sum_{n=1}^{\infty} n (1-p)^{n-1}$$

$$\sum_{n=1}^{\infty} n (1-p)^{n-1} = 1(1-p)^0 + 2(1-p) + 3(1-p)^2 + 4(1-p)^3 + \dots =$$

$$= 1 + 2q + 3q^2 + 4q^3 + \dots = \frac{1}{(1-q)^2}$$

$$E(N) = p \frac{1}{p^2} = \frac{1}{p} \quad (10)$$

Este resultado significa que el número promedio de pruebas hasta que ocurre el primer éxito es la inversa de la probabilidad  $p$  de ocurrencia del evento. Esto se asocia al concepto de **período de retorno**, sumamente utilizado en ingeniería.

$$\begin{aligned}
\text{Var}(N) &= E(N^2) - E^2(N) = \left[ \sum_{n=1}^{\infty} n^2 p (1-p)^{n-1} \right] - \left( \frac{1}{p} \right)^2 = \\
&= \frac{1}{p^2} \left[ \sum_{n=1}^{\infty} n^2 p^3 (1-p)^{n-1} - 1 \right] = \\
&= \frac{1}{p^2} \left[ \sum_{n=1}^{\infty} n^2 (1-q)^3 q^{n-1} - 1 \right] = \\
&= \frac{1}{p^2} \left[ \sum_{n=1}^{\infty} q^{n-1} (1-3q+3q^2+q^3) n^2 - 1 \right] \\
&\quad \sum_{n=1}^{\infty} q^{n-1} (1-3q+3q^2+q^3) n^2 = 1+q \\
\text{Var}(N) &= \frac{1}{p^2} (1+q-1)
\end{aligned}$$

$$\text{Var}(N) = \frac{1-p}{p^2} \quad (11)$$

El término período de retorno es muchas veces mal interpretado ya que se cree que pasarán 100 años entre ocurrencias de valores de una variable de interés superiores a un valor crítico; en realidad la probabilidad de que se produzca un valor mayor al crítico permanece invariable y en este caso igual a 0.01 para cualquier año, independientemente de lo que ocurrió en años anteriores, se está hablando de valor promedio.

### Ejemplo

Cuál sería la probabilidad de que una crecida con período de retorno 10 años ocurra antes o durante el décimo año en la ciudad de Santa Fe.

$$\text{siendo } T = 1/p \quad \rightarrow \quad p = 1/T = 1/10$$

$$\text{luego } P(X \leq 10) = 1 - (1 - 0.1)^{10} = 65\%$$

$$F(10) = 1 - (1 - p)^{10}$$

## Modelo Poisson

Considerando las condiciones que caracterizan al modelo Binomial, es necesario analizar que sucede con la distribución de la variable si el número de pruebas  $n$  se incrementa y la probabilidad  $p$  de éxito se hace cada vez más pequeña.

Si  $n$  se incrementa y  $p$  se hace pequeña, pero el número promedio de eventos en el intervalo total permanece constante e igual a  $\lambda = n \cdot p$  denominando a esta constante  $\lambda$  y considerando la función masa de probabilidad de  $x$  en el límite, esto es con  $p \rightarrow 0$  y  $n \rightarrow \infty$  y  $\lambda = n \cdot p$  a partir de la función del modelo Binomial:

$$\begin{aligned} \lambda = n p \quad \rightarrow \quad p &= \frac{\lambda}{n} \\ f(x) &= \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} = \\ &= \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \frac{n!}{(n-x)! n^x \left(1 - \frac{\lambda}{n}\right)^x} = \\ &= \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left[ \frac{n(n-1)(n-2)\dots(n-(x+1))(n-x)!}{(n-x)! \left(n \left(1 - \frac{\lambda}{n}\right)\right)^x} \right] \end{aligned}$$

La expresión entre corchetes tiene  $x$  términos en el numerador y  $x$  términos en el denominador, y para un  $n$  suficientemente grande, este término es igual a :

$$n^x / n^x = 1$$

con lo cual queda por evaluar la expresión  $\left(1 - \frac{\lambda}{n}\right)^n$  que cuando  $n \rightarrow \infty$  es igual a la constante  $e^{-\lambda}$ , esto permite expresar a la función de probabilidad del modelo de Poisson de la siguiente forma:

$$f(x) = P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad (12)$$

$\lambda$  parámetro del modelo.

Esta función deberá cumplir con la condición para ser una función de probabilidad:

$$\begin{aligned} \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} &= \\ &= e^{-\lambda} \left[ \frac{\lambda^0}{0!} + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots + \frac{\lambda^n}{n!} \right] \\ \left[ \frac{\lambda^0}{0!} + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots + \frac{\lambda^n}{n!} \right] &= e^{\lambda} \\ e^{-\lambda} \cdot e^{\lambda} &= e^{-\lambda+\lambda} = e^0 = 1 \end{aligned}$$

### Características

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x \frac{e^{-\lambda} \lambda^x}{x!} = \\ &= \lambda \sum_{x=1}^{\infty} x \frac{e^{-\lambda} \lambda^{x-1}}{x(x-1)!} \\ &= \sum_{x=1}^{\infty} \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!} = 1 \end{aligned}$$

$$E(X) = \lambda \quad (13)$$

$$\text{Var}(X) = E(X^2) - E^2(X)$$

$$\begin{aligned} E(X^2) &= \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda} \lambda^x}{x!} = \lambda \left[ \sum_{x=0}^{\infty} \frac{x^2 \lambda^{x-1} e^{-\lambda}}{x(x-1)!} \right] = \\ &= \lambda \left[ \sum_{x=1}^{\infty} \frac{x \lambda^{x-1} e^{-\lambda}}{(x-1)!} \right] \end{aligned}$$

si  $y = x - 1 \rightarrow y + 1 = x$ , luego

$$\lambda \left[ \sum_{y=0}^{\infty} (y+1) \frac{\lambda^y e^{-\lambda}}{y!} \right] = \lambda \left[ \sum_{y=0}^{\infty} y \frac{e^{-\lambda} \lambda^y}{y!} + \sum_{y=0}^{\infty} \frac{e^{-\lambda} \lambda^y}{y!} \right]$$

$$\sum_{y=0}^{\infty} y \frac{e^{-\lambda} \lambda^y}{y!} = E(Y) = \lambda$$

$$\sum_{y=0}^{\infty} \frac{e^{-\lambda} \lambda^y}{y!} = 1$$

luego

$$E(X^2) = \lambda(\lambda + 1) = \lambda^2 + \lambda$$

$$\text{Var}(X) = \lambda^2 + \lambda - \lambda^2$$

$$\text{Var}(X) = \lambda \quad (14)$$

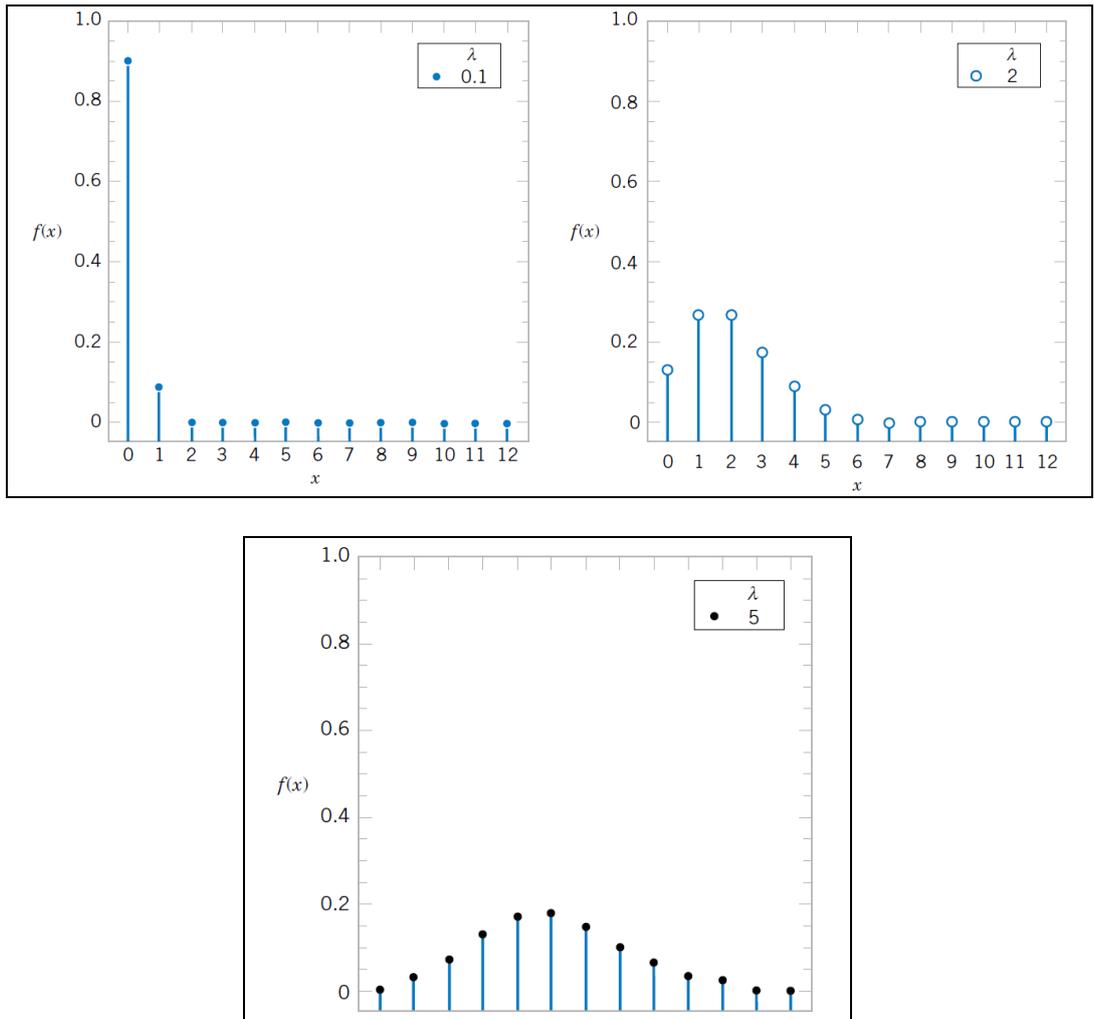
Generalmente este modelo se vincula a aquellos eventos que ocurren en una unidad de tiempo, luego el período de tiempo en el que se realiza el análisis constituye una secuencia de pruebas independientes cada una con distribución Binomial.

Si se tomara para el análisis un intervalo de tiempo igual al doble o al triple del inicial el parámetro es también igual al doble, al triple, etc., marcando esto la dependencia del tiempo de este modelo y por ello vinculado a los procesos estocásticos. Se entiende por procesos estocásticos a aquellos en los que interesa la secuencia en el tiempo de ocurrencia de eventos.

Dado un intervalo de números reales, asumiendo que ocurren al azar en todo el intervalo valores de una variable aleatoria; si el intervalo se puede dividir en subintervalos de pequeña longitud tal que:

- (1) la probabilidad de que más de una ocurrencia en un subintervalo es cero,
- (2) la probabilidad de una ocurrencia en un subintervalo es la misma para todos los subintervalos y proporcional a la longitud del subintervalo, y
- (3) el conteo en cada subintervalo es independiente de otros subintervalos,

Entonces el experimento aleatorio se denomina un *proceso de Poisson*.



**Figura N° 3** - Distribución de Poisson para distintos valores del parámetro

### *Ejemplo*

En cierta región se sabe que en un determinado mes del año ocurren, en promedio, 3 tormentas cuya precipitación supera los 150 mm. Determine la probabilidad de que en un determinado mes en esta zona ocurran menos de 5 tormentas de este tipo.

$$\lambda = 3$$

$$P(x < 5) = \sum_{x_i=0}^4 \frac{e^{-3} 3^{x_i}}{x_i!} = 0.8153$$

## Modelo Hipergeométrico

Este modelo surge cuando se realiza un muestreo sin reposición de una población finita y con sus elementos clasificados en dos categorías.

Si  $N$  es el total de elementos de los cuales hay  $k$  de una categoría y  $N-k$  de otra, al realizar una extracción de  $n$  elementos, *sin reposición*, cada extracción que se realice posteriormente es dependiente del resultado de la extracción anterior con lo cual va cambiando la probabilidad de éxito.

Para derivar la función correspondiente a la variable aleatoria  $x$ : número de éxitos o elementos pertenecientes a la categoría que se estudia, en una extracción, se deberán considerar todas las maneras posibles o combinaciones, de extraer  $x$  elementos de la categoría deseada, de los  $n$  extraídos y los restantes que pertenezcan a la otra categoría. El total de casos se obtiene de las combinaciones del total  $N$  extraídos de  $n$ . Luego la función de probabilidad para este modelo es:

$$P(X = x) = f(x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} \quad (15)$$

### Características

Para obtener las características es posible decir que la variable aleatoria  $X$  es la suma de  $n$  variables  $x_i$  como en el caso Binomial, pero con la diferencia que aquí las  $x_i$  son dependientes. Como para sumar las esperanzas no se necesita que las variables aleatorias sean independientes es posible obtener la esperanza de la siguiente forma:

$$E(X) = E(x_1) + E(x_2) + \dots + E(x_n)$$

donde en cada  $E(x_i)$  la probabilidad de  $x$  en la  $i$ ésima prueba es  $k/N$ , si no se sabe que ha ocurrido en pruebas anteriores o posteriores, luego:

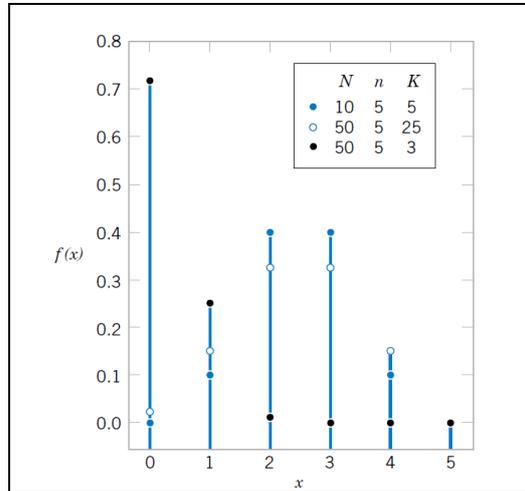
$$E(X) = n p = n \left( \frac{k}{N} \right) \quad (16)$$

La varianza:

$$Var(X) = n * p * q * \left( \frac{N-n}{N-1} \right) \quad (17)$$

siendo  $\left(\frac{N-n}{N-1}\right)$  el factor de corrección por muestreo sin reposición y población finita.

Cuando  $\left(\frac{n}{N}\right) \leq 0.05$  la distribución Hipergeométrica se aproxima a la Binomial.



**Figura N° 4** - Distribución Hipergeométrica para distintos valores de N, n y K

### **Ejemplo**

Un equipo incluye cinco ingenieros ambientales y nueve agrimensores. Si se eligen al azar cinco profesionales y se les asigna un proyecto, ¿cuál es la probabilidad de que el equipo del proyecto incluya exactamente a dos ingenieros ambientales?.

### **Solución:**

X: "número de ing. ambientales incluidos en el proyecto".

$$P(\text{Ing. Amb.}) = 5/14$$

$$P(\text{Agrim.}) = 9/14$$

$$P(X = 2) = \frac{\binom{5}{2} \binom{9}{3}}{\binom{14}{5}}$$

$$P(X = 2) = 0,42$$



Universidad Nacional del Litoral  
Facultad de Ingeniería y Ciencias Hídricas

## ESTADÍSTICA

**Ingenierías:** Recursos Hídricos-Ambiental-Agrimensura

---

## TEORÍA

*Mg. Ing. Susana Vanlesberg*  
Profesor Titular

# UNIDAD 4

## Modelos Probabilísticos

## Variable Continua

## MODELOS DE VARIABLE ALEATORIA CONTINUA

### Modelo exponencial

Este modelo surge al considerar el *tiempo hasta* la primera ocurrencia de un evento que pueda ser considerado como proceso de Poisson.

Si la variable aleatoria es ahora el tiempo transcurrido hasta que se verifica la primera ocurrencia, entonces será una variable continua, la probabilidad que T exceda algún valor t es lo mismo que decir que **no** se verificaron ocurrencias en ese intervalo de longitud t, lo que es equivalente a decir que la variable aleatoria **Nº de ocurrencias** de tipo Poisson toma el valor 0:

$$P(T > t) = P(x = 0)$$
$$P(x = 0) = \frac{e^{-\lambda t} \lambda^0}{0!} = e^{-\lambda t}$$

Esto permite obtener la función de distribución de T variable continua:

$$F(t) = P(T \leq t) = 1 - P(T > t) = 1 - e^{-\lambda t} \quad t \geq 0 \quad (1)$$

La función de densidad se obtiene derivando la expresión anterior:

$$f(t) = \lambda e^{-\lambda t} \quad (2)$$

Cumpliendo con las propiedades de estacionariedad e independencia de los procesos de Poisson,  $e^{-\lambda t}$  da la probabilidad de que no se verifiquen eventos en algún intervalo de tiempo de longitud t, estando o no el origen en el tiempo 0.

Por ejemplo si el origen estuviese en el momento de la enésima ocurrencia,  $e^{-\lambda t}$  representaría la probabilidad de que el tiempo hasta la próxima ocurrencia (n+1) sea mayor que t.

Esto quiere demostrar sencillamente que los tiempos entre ocurrencias de eventos de tipo Poisson son independientes y se distribuyen en forma exponencial.

## Características

$$E(T) = \int_0^{\infty} t \lambda e^{-\lambda t} dt$$

$$\text{si } u = t\lambda \quad du = \lambda dt \quad \frac{du}{\lambda} = dt$$

$$\frac{1}{\lambda} \int_0^{\infty} u e^{-u} du = \frac{1}{\lambda} [e^{-u}(-u-1)]_0^{\infty}$$

$$E(T) = \frac{1}{\lambda} \quad (3)$$

Recordar que  $\lambda$  en los procesos de Poisson, representa el número promedio de ocurrencias, aquí,  $1/\lambda$  representa el tiempo promedio entre ocurrencias.

$$\text{Var}(T) = E(T^2) - E^2(T)$$

$$E(T^2) = \int_0^{\infty} t^2 \lambda e^{-\lambda t} dt$$

$$\text{si } v = t^2 \quad dv = 2t dt$$

$$\lambda e^{-\lambda t} dt = du \quad u = \int_0^{\infty} \lambda e^{-\lambda t} dt = -e^{-\lambda t}$$

$$E(T^2) = [-e^{-\lambda t} t^2]_0^{\infty} - \int_0^{\infty} -e^{-\lambda t} 2t dt =$$

$$= 0 + 2 \int_0^{\infty} e^{-\lambda t} t dt \frac{\lambda}{\lambda} = \frac{2}{\lambda} \left( \frac{1}{\lambda} \right)$$

$$E(T^2) = \frac{2}{\lambda^2}$$

$$\text{Var}(T) = \frac{2}{\lambda^2} - \left( \frac{1}{\lambda} \right)^2 = \frac{1}{\lambda^2} \quad (4)$$

Una característica de los procesos de Poisson es que no tienen memoria. Esto significa que el comportamiento futuro es independiente de lo registrado en el presente o en el pasado.

Para comprenderlo mejor se puede calcular la probabilidad condicional de T dado que  $T > t_0$ ; o sea la probabilidad condicional del tiempo entre ocurrencias dado que no se han registrado ocurrencias antes del tiempo  $t_0$ .

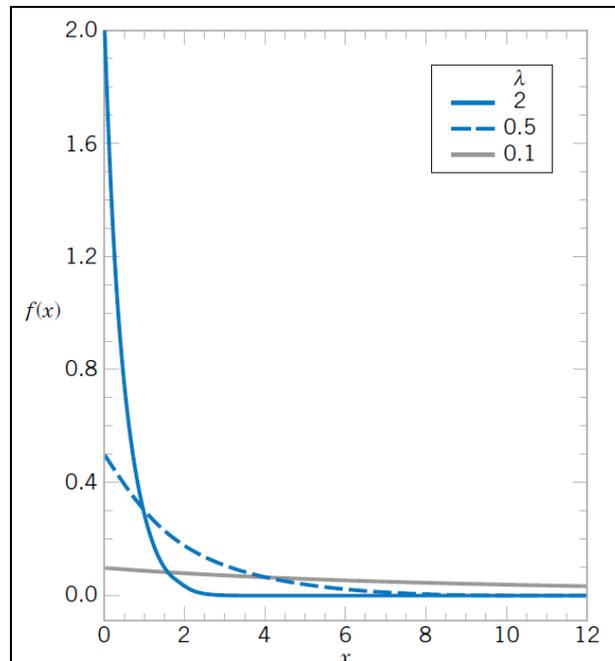
$$P(T \leq t / T \geq t_0) = \frac{P(T \leq t \cap T \geq t_0)}{P(T \geq t_0)}$$

*parat <math>t\_0</math> el numeradores cero*  
*parat  $\geq t_0$  es iguala  $P(t_0 \leq T \leq t)$*

$$P(T \leq t / T \geq t_0) = \frac{P(t_0 \leq T \leq t)}{P(T \geq t_0)} = \frac{F(t) - F(t_0)}{1 - F(t_0)} =$$

$$= \frac{1 - e^{-\lambda t} - (1 - e^{-\lambda t_0})}{1 - (1 - e^{-\lambda t_0})} = 1 - e^{-\lambda(t-t_0)} = 1 - e^{-\lambda T}$$

la expresión obtenida coincide con la del modelo (4), demostrando así que hay independencia respecto a ocurrencias pasadas.



**Figura N° 1-** Distribución exponencial para distintos valores del parámetro

**Ejemplo**

Si ocurrieran tres lluvias cuyas magnitudes causaran daños sumamente importantes, de duración 1 hora y cada 10 años; se quiere determinar la probabilidad de que transcurra menos de un año hasta que ocurra la próxima:

$$P(T \leq 1) = 1 - e^{(-0.3 * 1)} = 0.26$$

## Modelo Gamma

Este modelo surge al considerar como variable el tiempo transcurrido hasta que se verifique la **k-ésima ocurrencia** de un proceso de tipo Poisson. Cada uno de los tiempos entre ocurrencias son independientes y exponencialmente distribuidos con parámetro común  $\lambda$ . La variable  $X_k$ , es la suma de esos tiempos  $T_1 + T_2 + \dots + T_k$

Esto origina la expresión para la función de densidad del modelo:

$$f(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)} \quad \text{con } x > 0 \quad (5)$$

Se dice que  $x$  está distribuida como Gamma con parámetros  $\lambda$  y  $k$ , con  $(k-1)! = \Gamma(k)$  (para valores de  $k$  enteros). Esta función se denomina función Gamma, que para cualquier valor de  $k$  se obtiene a través de la siguiente expresión:

$$\Gamma(k) = \int_0^{\infty} x^{k-1} e^{-x} dx \quad \text{con } r > 0 \quad (6)$$

Si  $k$  no es entero, pueden obtenerse los valores de  $\Gamma(k)$  en tablas, o a través de aproximaciones numéricas, o en softwares.

La función de densidad debe cumplir con las propiedades de una función de probabilidad:

$$\begin{aligned} f(x) &= \frac{\int_0^{\infty} \lambda (\lambda x)^{(k-1)} e^{-\lambda x} dx}{\int_0^{\infty} x^{(k-1)} e^{-x} dx} = \\ &= \frac{\lambda^k \int_0^{\infty} x^{(k-1)} e^{-\lambda x} dx}{\int_0^{\infty} x^{(k-1)} e^{-x} dx} = 1 \end{aligned}$$

La función acumulativa se obtiene integrando la anterior:

$$F(x) = \int_0^x f(t) dt = \frac{\Gamma(k, \lambda x)}{\Gamma(k)} = \frac{\text{función gamma incompleta}}{\text{función gamma}} \quad (7)$$

Ya que la distribución gamma se ha originado como la suma de  $k$  variables independientes idénticamente distribuidas como exponenciales, sus características pueden derivarse de ese hecho:

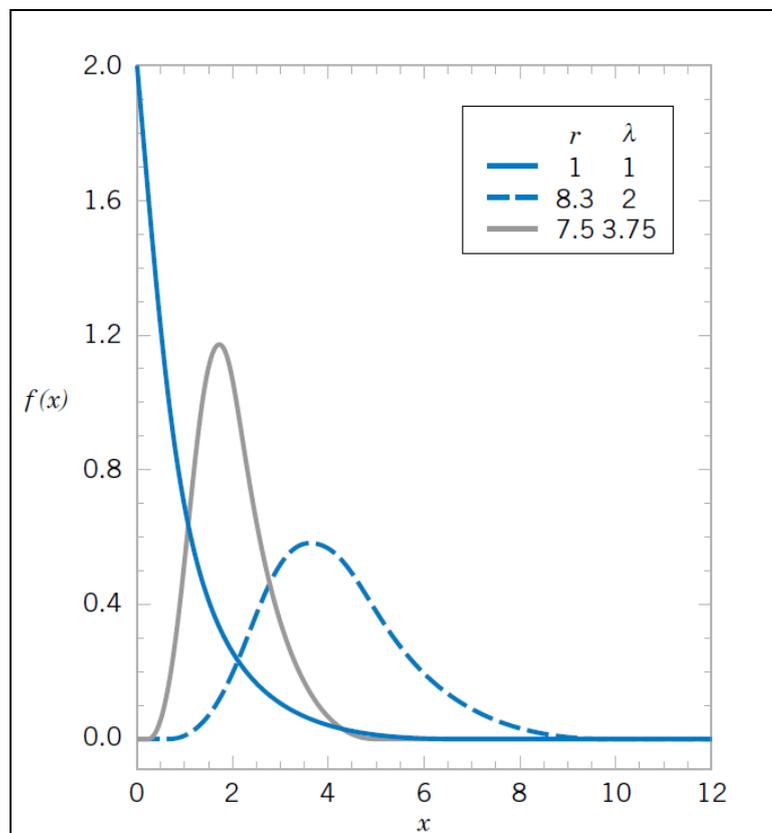
$$E(X) = \frac{k}{\lambda} \quad (8)$$

$$\text{Var}( X ) = \frac{k}{\lambda^2} \quad (9)$$

La forma de la función de densidad de este modelo indica el porqué de su amplia aplicación y, sobre todo, en las ciencias de la tierra donde muchas variables están limitadas a valores positivos y presentan asimetría hacia la derecha.

$\lambda$  puede interpretarse como un factor de escala, y  $k$  como factor de forma.

Esta distribución ha sido ampliamente utilizada para describir fenómenos tales como caudales máximos. También se la conoce como *distribución Pearson III*.



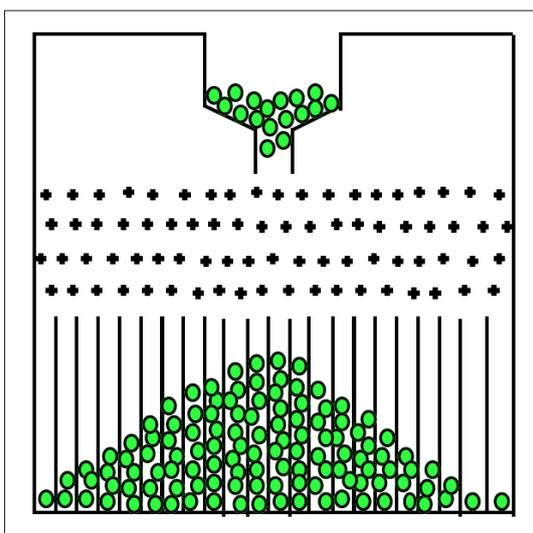
**Figura N° 2** - Función de densidad Gamma para distintos valores de los parámetros

## MODELO NORMAL

La distribución Normal es la distribución continua más importante del Cálculo de Probabilidades y de la Estadística. Aparece por primera vez en 1733 en los trabajos de DE MOIVRE relativos al cálculo de la distribución límite de una variable binomial.

Posteriormente, en 1809, GAUSS y más tarde, en 1812, LAPLACE la estudiaron en relación con la teoría de errores de datos experimentales, al tratar de hallar el valor correcto más probable entre una serie de medidas.

Una primera aproximación de la distribución Normal puede observarse con el experimento que realizó SIR FRANCIS GALTON, que construyó un ingenioso aparato, formado por un tablero inclinado, en el que se distribuyen regularmente un sistema de clavos, para acabar finalmente en compartimentos estrechos. Al deslizar muchas bolas desde un depósito superior, estas chocan con los clavos, y se alejan más o menos de la línea central de caída. Las alturas alcanzadas por las bolas en los compartimentos estrechos da una idea de la curva de la distribución Normal.



El nombre de distribución Normal se debe al hecho de que una mayoría de las variables aleatorias de la Naturaleza siguen esta distribución, lo que hizo pensar que todas las variables continuas de la Naturaleza eran normales, llamando a las demás distribuciones "anormales". No obstante, hoy en día, ya no se piensa de la misma manera, ya que ningún estadístico dice que una distribución que no sea normal, es anormal. La distribución normal es la más importante por sus propiedades sencillas, porque aparece frecuentemente en la Naturaleza, (fenómenos relacionados con psicología, biología, etc.), y por una propiedad de algunos fenómenos que se aproximan asintóticamente a la distribución Normal (Teorema Central del Límite).

Este modelo suele conocerse como **Modelo de las sumas**, ya que la incertidumbre en algunas variables puede ser el resultado de efectos combinados de algunas causas que contribuyen, siendo difícil de separar y observar a cada una. En algunas situaciones si

se conoce el mecanismo por el cual las causas individuales afectan a la variable de interés se puede determinar un modelo para la variable resultante sin estudiar en detalle los efectos individuales, particularmente no es necesario conocer la distribución de las causas.

El modelo de la variable resultado, presenta una función de densidad doble exponencial cuya forma es la de una campana. Este hecho está contenido en el Teorema del Límite Central, el cual es uno de los resultados más importantes de la Teoría de Probabilidad. Este teorema será desarrollado más adelante pero se puede hacer referencia a su enunciado: *bajo condiciones generales, cuando el número de variables que intervienen en una suma que origina una variable aleatoria, es cada vez más grande, la distribución de esta suma tiende a aproximarse al modelo Normal.*

La inmensa importancia práctica del modelo Normal reside entre otras razones en que lo que se plantea en el teorema del Límite Central, puede hacerse sin el conocimiento exacto de:

- las distribuciones marginales de cada variable que interviene en la suma, - de su número, - de su distribución conjunta. Ya que la variación aleatoria en algunos fenómenos naturales se origina de un número de variaciones aditivas, no debe sorprender el hecho que gráficos que pueden aproximarse a este modelo, se observen con frecuencia en la naturaleza.

La función de densidad que caracteriza esta distribución es una doble exponencial con la siguiente forma:

$$f(x) = ke^{-c(x-m)^2}, -\infty \leq x \leq \infty \quad (10)$$

siendo **m** la distancia al centro de la distribución que por la simetría de la distribución es igual a su media  $\mu$ .

Las constantes k y c se determinan considerando que f(x) debe ser una función de probabilidad y también se utiliza la expresión de la varianza.

$$\int_{-\infty}^{\infty} k e^{-c(x-\mu)^2} dx = 1$$

$$k \int_{-\infty}^{\infty} e^{-c(x-\mu)^2} dx = 1$$

llamando  $(x - \mu) = y$ ,  $dx = dy$ , luego

$$k \int_{-\infty}^{\infty} e^{-c y^2} dy = 1$$

$$\text{siendo } \int_0^{\infty} e^{-a^2 x^2} dx = \frac{\sqrt{\pi}}{2a} \quad a^2 = c \longrightarrow a = \sqrt{c}$$

por simetría se tiene:

$$k 2 \frac{\sqrt{\pi}}{2\sqrt{c}} = 1$$

$$k = \frac{\sqrt{c}}{\sqrt{\pi}}$$

$$\text{Var}(x) = \frac{\sqrt{c}}{\sqrt{\pi}} \int_{-\infty}^{\infty} (X - \mu)^2 e^{-c(X-\mu)^2} dx$$

Para resolver esta integral se usa un resultado proveniente de la resolución de integrales impropias dependientes de un parámetro:

$$\int_0^{\infty} e^{-c x^2} dx = \frac{1}{2} \frac{\sqrt{\pi}}{\sqrt{c}} \text{ para } c > 0$$

$$\text{llamando } F(c) = \int_0^{\infty} e^{-c x^2} dx$$

y aplicando la siguiente propiedad

$$F(x) = \int_a^{\infty} f(t, x) dt \longrightarrow F'(x) = \int_a^{\infty} \frac{\partial f}{\partial x}(t, x) dt$$

en este caso origina:

$$f(x, c) = e^{-c x^2} \quad \text{luego } \frac{\partial f}{\partial c} = -x^2 e^{-c x^2}$$

$$\text{con lo cual } F'(c) = \int_0^{\infty} -x^2 e^{-c x^2} dx$$

$$\text{pero como } F'(c) = \frac{dF}{dc} = \frac{d}{dc} \left( \frac{1}{2} \frac{\sqrt{\pi}}{\sqrt{c}} \right) = -\frac{\sqrt{\pi}}{4} c^{-\frac{3}{2}}$$

resulta:

$$\int_0^{\infty} -x^2 e^{-cx^2} dx = -\frac{\sqrt{\pi}}{4} c^{-\frac{3}{2}}$$

$$\text{luego } \int_{-\infty}^{\infty} x^2 e^{-cx^2} dx = \frac{\sqrt{\pi}}{2} c^{-\frac{3}{2}}$$

$$\text{y } \frac{\sqrt{c}}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^2 e^{-cx^2} dx = \frac{\sqrt{c}}{\sqrt{\pi}} \frac{\sqrt{\pi}}{2} c^{-\frac{3}{2}} = \frac{c^{-1}}{2}$$

$$\text{Var}(X) = \frac{1}{2c} \text{ de donde}$$

$$c = \frac{1}{2\sigma^2}$$

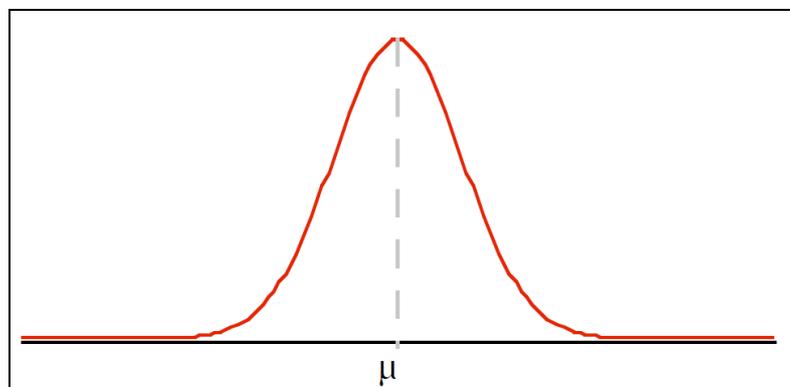
Por lo cual la expresión de la función de densidad es la siguiente:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty \leq x \leq \infty \quad (11)$$

La función acumulativa:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt \quad (12)$$

Siendo  $\mu$  y  $\sigma$  sus parámetros.  $\sigma > 0$  desvío estándar y  $-\infty < \mu < \infty$  la media. La función queda completamente definida si se conocen estos dos valores.



**Figura N° 3 – Distribución Normal**

Análisis del significado de los términos de la función:

$-\left[\frac{(x-\mu)}{\sigma}\right]^2$ : es la parte práctica de la función ya que contiene un valor determinado de la variable y a los parámetros. El hecho de estar elevado al cuadrado hace que dos valores distintos de la variable, que tengan la misma desviación absoluta de la media  $\mu$ , van a tener igual valor de la función de densidad. Esto muestra claramente la característica de simétrica alrededor de  $\mu$ , que posee la distribución Normal.

-Al ser este exponente negativo, cuanto mayor es la desviación de  $x$  respecto de  $\mu$ , será menor la densidad de probabilidad de  $x$ , esto es lo que se observa en las colas de la distribución. Ahora cuando  $x$  coincide con  $\mu$ , el exponente es cero y el valor de la función de densidad es máxima en ese punto e igual a  $1/\sigma \pi^{1/2}$ , o sea que  $x=\mu$  es el valor del modo.

- Presenta puntos de inflexión en los puntos de abscisas  $\mu + \sigma$  y  $\mu - \sigma$ , donde cambia de concavidad (lo que determina que cuanto mayor sea  $\sigma$ , más achatada sea la curva).

El punto de inflexión se obtiene al igualar a cero la derivada segunda, por lo tanto:

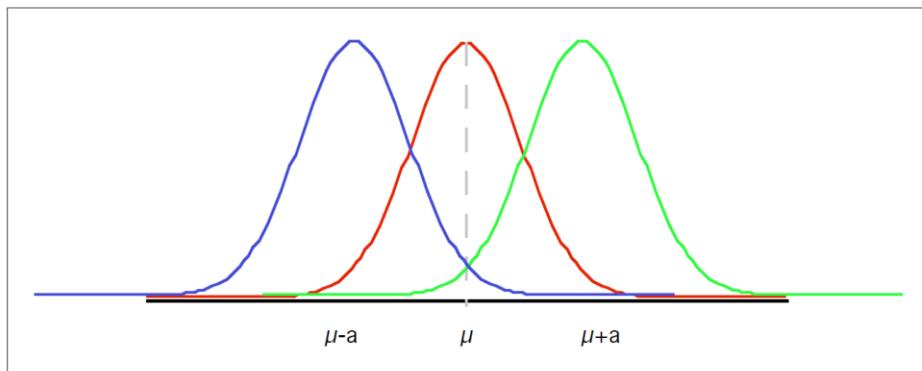
$$f''(x) = 0 \Leftrightarrow 1 - \frac{(x-\mu)^2}{\sigma^2} = 0 \Leftrightarrow \frac{(x-\mu)}{\sigma} = \pm 1 \Rightarrow x = \mu \pm \sigma$$

- Es asintótica al eje de abscisas, pues  $e^x$  tiende a 0 cuando  $x$  tiende a infinito, entonces:

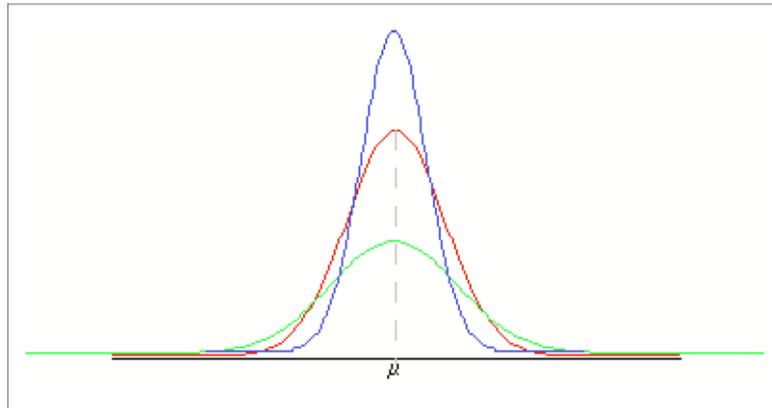
$$\lim_{x \rightarrow +\infty} f(x) = 0$$

es decir, el eje OX es una asíntota horizontal, e igual para  $x$  tendiendo a  $-\infty$ .

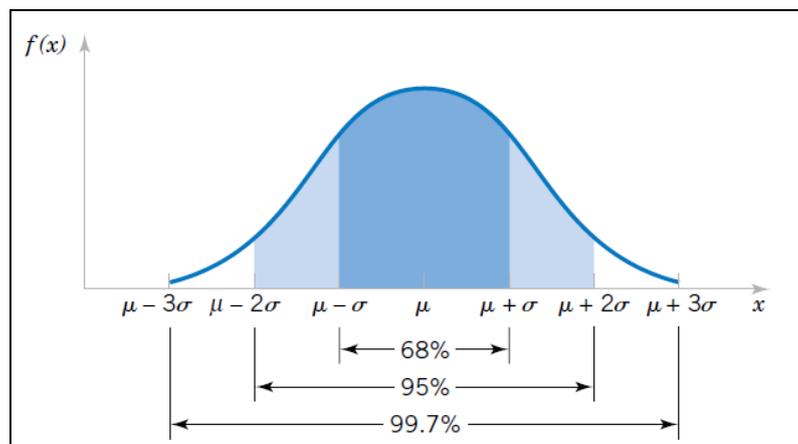
-El valor de  $\mu$  hace desplazar al modelo hacia la izquierda o la derecha, mientras que el valor de  $\sigma$  cambia su forma sin desplazarlo. Este efecto se observa en las siguientes figuras:



**Figura N° 4** – Efecto de variar  $\mu$  con  $\sigma$  fijo



**Figura N° 5 – Efecto de variar  $\sigma$  con  $\mu$  fijo**



**Figura N° 6 - Áreas en la distribución Normal**

### Características

Se obtendrán ahora el valor de la esperanza y de la varianza:

$$\begin{aligned}
 E(X) &= \int_{-\infty}^{\infty} x f(x) dx = \\
 &= \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} dx = \\
 \text{siendo } \frac{x-\mu}{\sigma} &= y \quad x = \sigma y + \mu \quad dx = \sigma dy \\
 E(X) &= \int_{-\infty}^{\infty} (\sigma y + \mu) \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{1}{2} y^2} \sigma dy =
 \end{aligned}$$

Integrando por partes, de la siguiente manera:

$$\begin{aligned}
&= \frac{\sigma}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} (\sigma y + \mu) e^{-\frac{y^2}{2}} dy = \\
&= \frac{1}{\sqrt{2\pi}} \left[ \int_{-\infty}^{\infty} (\sigma y) e^{-\frac{y^2}{2}} dy + \int_{-\infty}^{\infty} \mu e^{-\frac{y^2}{2}} dy \right] = \\
&= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\frac{y^2}{2}} dy + \mu \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \\
&= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} d\left(\frac{y^2}{2}\right) + \mu = \\
&= \frac{\sigma}{\sqrt{2\pi}} \left[ -e^{-\frac{y^2}{2}} \right]_{-\infty}^{\infty} + \mu = 0 + \mu \\
&E(X) = \mu
\end{aligned} \tag{13}$$

$$\begin{aligned}
Var(X) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x - \mu)^2 e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\
\frac{x - \mu}{\sigma} &= y \quad dy = \frac{dx}{\sigma} \\
\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma^2 y^2 e^{-\frac{1}{2}y^2} \sigma dy &= \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\frac{1}{2}y^2} d\left(\frac{y^2}{2}\right) \\
u = y \quad dv &= e^{-\frac{1}{2}y^2} d\left(\frac{y^2}{2}\right) \Rightarrow v = -e^{-\frac{y^2}{2}}
\end{aligned} \tag{14}$$

$$\begin{aligned}
Entonces, \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy &= \sigma^2 \left\{ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2} dy \right\} \\
Var(X) &= \sigma^2
\end{aligned}$$

Otras características del modelo Normal que merecen destacarse:

- Como ya se mencionó al principio, muchas variables aleatorias continuas se distribuyen en forma aproximadamente Normal. Es de destacar que los errores de mediciones repetidas de alguna dimensión particular, se dice que tienen distribución Normal. Esto se debe a que cualquier medida se supone que es igual a un valor verdadero más un error; este error puede considerarse como el resultado de un gran conjunto de factores que están presentes en ese momento, y cada factor ejerce un

pequeño efecto en la magnitud y sentido del error. Los errores actúan independientemente y con igual fuerza para aumentar o disminuir el valor de la medición observada, y a largo plazo se anulan. Esto hace considerar a los errores de medida como distribuidos Normalmente, con valor medio cero, y se los suele denominar "errores al azar".

- Este modelo sirve como buena aproximación de modelos discretos, como el Binomial o el de Poisson, bajo circunstancias especiales.

- El supuesto de Normalidad de las poblaciones permite obtener buenos resultados de métodos elaborados bajo este supuesto, aunque en realidad no se cumpla en forma estricta.

Muchos estadísticos calculados a partir de grandes muestras se distribuyen en forma aproximadamente Normal, lo cual facilita el trabajo de inferencia estadística.

### Modelo Normal Estándar

Se denomina modelo estándar a aquel cuya media es 0 y su desvío es uno. La variable  $z=(x-\mu)/\sigma$  se denomina *variable aleatoria estandarizada*, cuya media es 0 y su desvío es unitario.

Las funciones del modelo son:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad -\infty < z < +\infty \quad (15)$$

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{1}{2}t^2} dt \quad (16)$$

La distribución Normal se encuentra tabulada y existen, además, rutinas computacionales que permiten trabajar con ella.

Para fines prácticos y para ganar en eficiencia y rapidez, el concepto de distribución estándar es muy importante.

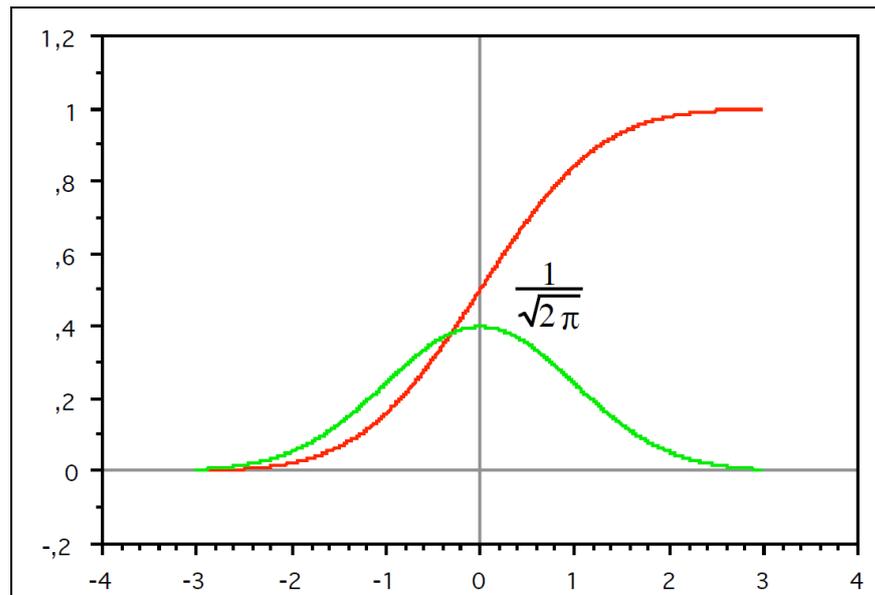
Las características de este modelo son las siguientes:

$$E\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma} E(x-\mu) = \frac{1}{\sigma} [E(x) - E(\mu)] = \frac{1}{\sigma} (\mu - \mu)$$

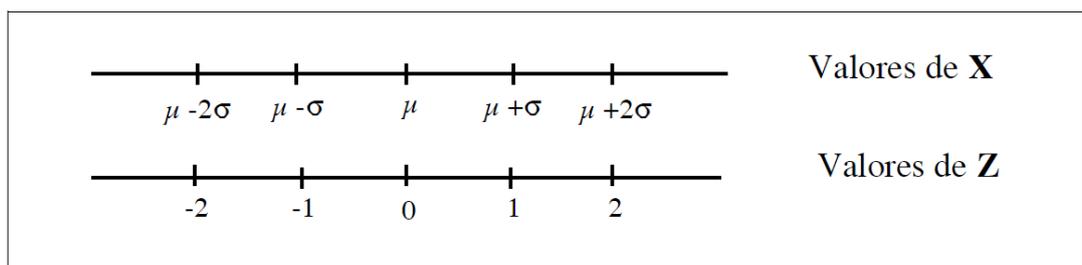
$$E(z) = 0 \quad (17)$$

$$\begin{aligned} \text{Var}(z) &= \text{Var}\left(\frac{x - \mu}{\sigma}\right) = \frac{1}{\sigma^2} \text{Var}(x - \mu) = \frac{1}{\sigma^2} \text{Var}(x) - 0 = \frac{\sigma^2}{\sigma^2} \\ \text{Var}(\mu) &= E[\mu - E(\mu)]^2 = E(\mu - \mu)^2 = 0 \end{aligned} \quad (18)$$

$$\text{Var}(z) = 1$$



**Figura N° 7** – Funciones del Modelo Normal Estándar



La simetría del modelo Normal respecto a su media implica que los momentos centrados de orden impar sean cero. Los de orden par pueden considerarse a partir de la siguiente expresión:

$$\mu_n = E(x - \mu)^n = \frac{n!}{2^{\frac{n}{2}} \left(\frac{n}{2}\right)!} \sigma^n \quad n = 2, 4, \dots \quad (19)$$

La asimetría será, por lo tanto, igual a cero:

$$As = \frac{\mu_3}{\sigma^3} = 0 \quad (20)$$

y la kurtosis:

$$K = \frac{\mu_4}{\sigma^4}$$
$$\mu_4 = \frac{4!}{2^2 \left(\frac{4}{2}\right)!} \sigma^4 = 3\sigma^4 \quad (21)$$

$$K = 3$$

Es por esta característica que suelen compararse los coeficientes de kurtosis de distintos modelos con el de la Normal, permitiendo clasificarlos según sean mayores, menores o iguales que 3.

### **Distribución de la suma de variables aleatorias Normales**

Como se ha expresado anteriormente, el Teorema del Límite Central (que más adelante se expondrá con detalle), expresa que la suma de un número dado de variables aleatorias tiende a distribuirse en forma Normal. Es de esperar que si se suma un número de variables aleatorias independientes y Normales cada una, la suma será distribuida en forma Normal.

$$x_1 \sim N(\mu_1, \sigma_1)$$

$$x_2 \sim N(\mu_2, \sigma_2)$$

·

·

·

$$x_n \sim N(\mu_n, \sigma_n)$$

$$U = \sum_{i=1}^n x_i$$

con parámetros

$$E(U) = E\left(\sum_{i=1}^n x_i\right) = \sum_{i=1}^n E(x_i) = \sum_{i=1}^n \mu_i$$

$$\sigma^2(U) = \sigma^2\left(\sum_{i=1}^n x_i\right) = \sum_{i=1}^n \sigma^2(x_i) = \sum_{i=1}^n \sigma_i^2$$

$$\text{entonces } U \sim N\left(\sum \mu_i; \sqrt{\sum \sigma_i^2}\right)$$

(22)

El uso de la distribución Normal es generalizado en muchas ciencias, especialmente cuando se agregan valores para obtener valores anuales o mensuales.

### Aproximación del modelo Binomial al Normal

Si se considera una variable aleatoria con distribución Binomial, cuando n aumenta y p no varía, la variable aleatoria

$$\frac{x - np}{\sqrt{npq}}$$

se distribuye aproximadamente Normal (0,1):

$$P(X \leq x_0) = F_N\left(\frac{x_0 - np}{\sqrt{npq}}\right) \quad (23)$$

siendo  $x_0$ , un valor cualquiera de x.

En este caso hay que tener en cuenta que  $X$  era una variable aleatoria discreta y queremos tratarla como continua, por lo que es preciso hacer una corrección para continuidad. Así se verifica que:

$$P(\mathbf{X} = 3) = P(2.5 < \mathbf{X} \leq 3.5)$$

$$P(\mathbf{X} \leq 3) = P(\mathbf{X} \leq 3.5)$$

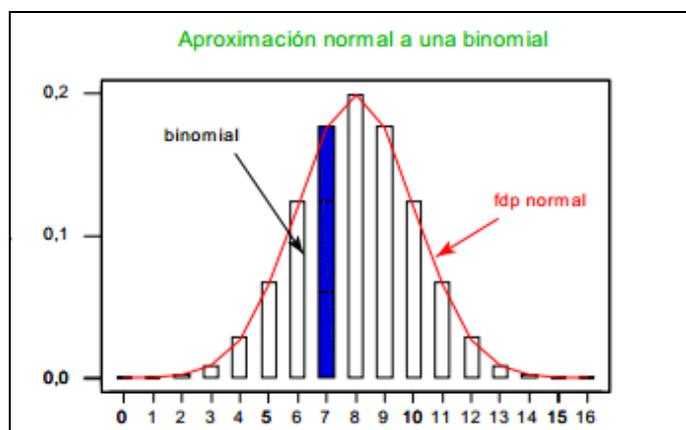
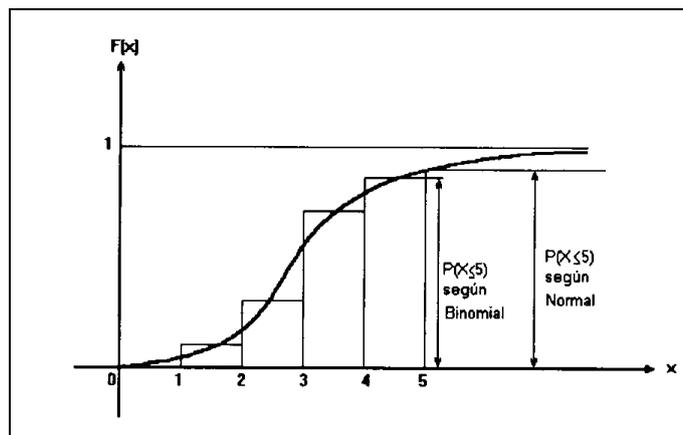
$$P(\mathbf{X} < 3) = P(\mathbf{X} \leq 2.5)$$

$$P(X \leq x_0) = F_N \left( \frac{x_0 + \frac{1}{2} - np}{\sqrt{npq}} \right)$$

El primer miembro de la última ecuación pertenece a la distribución Binomial, cuya función de distribución es escalonada. El segundo miembro pertenece a la distribución Normal que es continua. A pesar de esta diferencia, para grandes valores de  $n$  ambas pueden considerarse iguales para cada valor, ya que la función continua pasa por la mitad de cada escalón de la función discreta. Esto es una buena aproximación para valores de  $n \cdot p > 5$ , cuando  $p \leq 1/2$  y  $n \cdot q > 5$  con  $p > 1/2$ .

Obviamente éstas no son igualdades ciertas, pero permiten tratar la variable discreta como continua.

Si en lugar de trabajar con una variable aleatoria binomial partiésemos de una variable de Poisson, la aproximación sería absolutamente similar.



**Figura N° 8 – Aproximación Binomial con la Normal**

## DISTRIBUCIONES RELACIONADAS CON LA DISTRIBUCIÓN NORMAL

### Distribución Chi-cuadrado

Este modelo describe la distribución de la suma de los cuadrados de  $v$  variables aleatorias independientes, con distribución  $N(0,1)$ :

$$V = \sum_{i=1}^v X_i^2 \sim \chi_v^2 \quad (24)$$

Si las variables que intervienen en la suma no fuesen  $N(0,1)$ , se las debería estandarizar, con lo cual se obtendría:

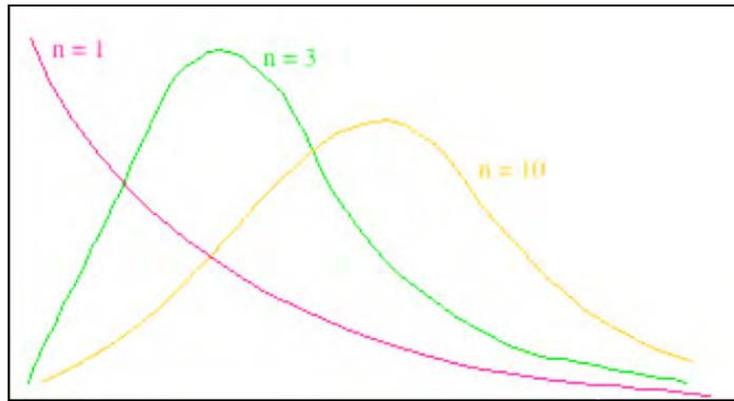
$$V = \sum_{i=1}^v \left( \frac{X_i - \mu_i}{\sigma_i} \right)^2 \sim \chi_v^2 \quad (25)$$

$V$  es también una variable aleatoria, ya que es el resultado de la suma de variables aleatorias. Por ser una suma de cuadrados, varía desde 0 a infinito.

La función de densidad surge de considerar la función  $N(0,1)$  de cada componente al cuadrado, y es la siguiente:

$$f(\chi^2) = \begin{cases} \frac{1}{2^{\frac{v}{2}} \Gamma\left(\frac{v}{2}\right)} \chi^{2\left(\frac{v-2}{2}\right)} e^{-\frac{\chi^2}{2}}, & \text{para } \chi^2 > 0 \\ 0, & \text{en otro caso} \end{cases} \quad (26)$$

siendo  $\Gamma$  la función gamma y  $v$  el número de variables al cuadrado que intervienen en la suma, y se denominan "grados de libertad".



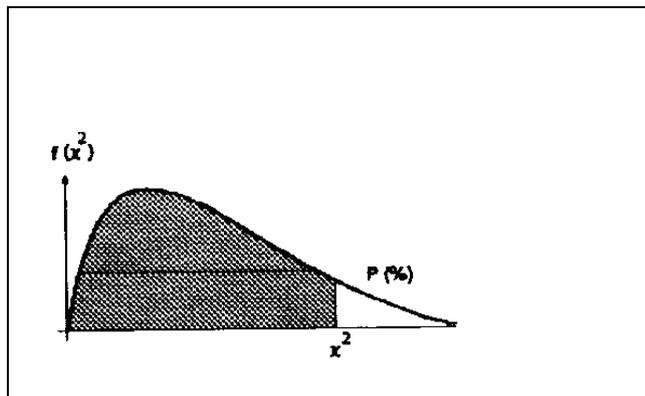
**Figura N° 9 – Distribución chi- cuadrado**

Las características de esta distribución son:

$$E(\xi^2)=v \quad (27)$$

$$\sigma^2(\xi^2)=2v$$

Las tablas permiten evaluar probabilidades del siguiente tipo:



### Propiedades de la distribución chi-cuadrado

- 1.- La variable solo puede tomar valores positivos.
- 2.- Es asimétrica.
- 3.- Depende del parámetro  $\nu$  (grados de libertad).
- 4.- Su esperanza matemática es  $\nu$ , y su varianza,  $2\nu$ .
- 5.- *Propiedad aditiva o reproductiva* :Si  $\chi^{2n}$  y  $\chi^{2m}$  son dos variables Chi cuadrado con  $n$  y  $m$  grados de libertad respectivamente, independientes entre sí, entonces la suma de las dos variables es una variable Chi-cuadrado con  $n+m$  grados de libertad. Esto se puede generalizar a la suma de cualquier número de variables Chi-cuadrado, independientes.
- 6.- Al aumentar el número de grados de libertad, la distribución Chicuadrado se aproxima asintóticamente a una distribución normal.

## Distribución t de Student

Este modelo surge como el cociente entre una variable  $N(0,1)$  y la raíz cuadrada de una variable aleatoria distribuida  $\chi^2$  dividida por sus grados de libertad, siendo estas dos variables independientes:

$$t_\nu = \frac{x}{\sqrt{\frac{V}{\nu}}}, \quad \begin{cases} x \sim N(0,1) \\ V \sim \chi_\nu^2 \end{cases} \quad (28)$$

El rango de la variable t varía entre  $-\infty$  y  $+\infty$ , y su función de densidad surge de las distribuciones de las variables aleatorias componentes:

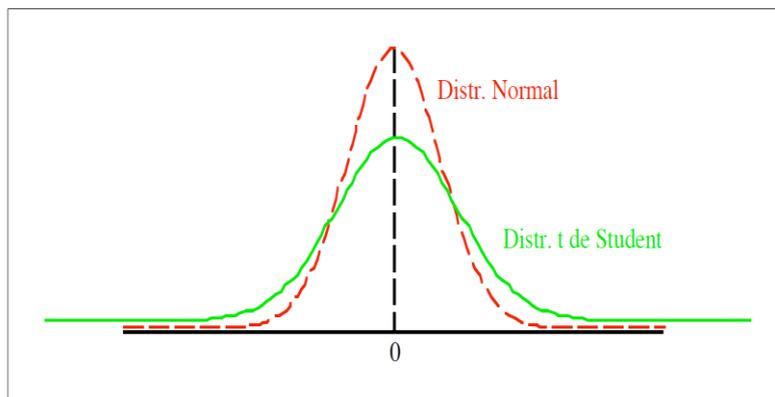
$$f(t) = \frac{1}{\sqrt{\pi\nu}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad \text{para } -\infty < t < \infty \quad (29)$$

siendo  $\nu$  los grados de libertad asociados a la componente  $\chi^2$

Sus características son:

$$\begin{aligned} E(t) &= 0 \\ \sigma^2(t) &= \frac{\nu}{\nu-2}, \quad \nu > 2 \end{aligned} \quad (30)$$

La gráfica es parecida a la de la distribución Normal, y es simétrica respecto del cero. Para grandes valores de  $\nu$  se la puede aproximar a la Normal.



**Figura N° 10** - Distribución de Student

Existen tablas que brindan probabilidades para distintos valores de la variable y algoritmos computacionales que permiten obtener probabilidades con igual o mayor grado de precisión que las tablas.

**Propiedades de la distribución "t"**

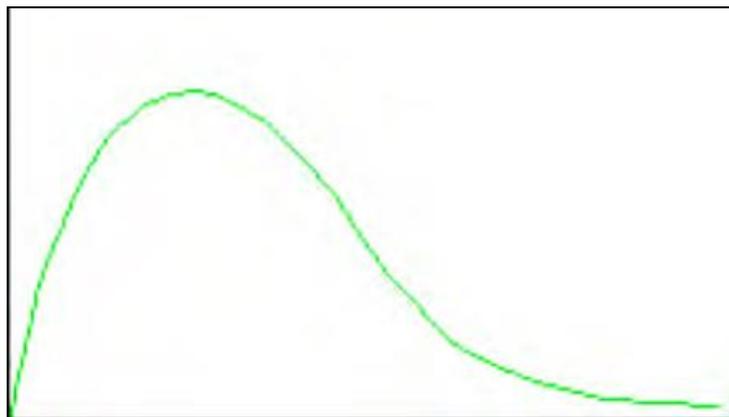
- 1.- Depende de un único parámetro, el número de grados de libertad.
- 2.- El rango de la variable es todo el eje real  $(-\infty, +\infty)$ .
- 3.- Su gráfica es simétrica respecto al eje de ordenadas OY.
- 4.- El valor  $x = 0$  es la media, mediana y moda de la distribución.
- 5.- Al aumentar  $n$ , se va haciendo cada vez más apuntada la gráfica de su función de densidad, siendo el límite para  $n \rightarrow \infty$  la curva normal tipificada.

**Distribución F de Snedecor**

Si X e Y son dos variables aleatorias independientes cada una con distribución  $\chi^2$ , con  $\nu_1$  y  $\nu_2$  grados de libertad; luego la distribución F está definida por el siguiente cociente:

$$F_{\nu_1, \nu_2} = \frac{\frac{X}{\nu_1}}{\frac{Y}{\nu_2}} = \frac{\chi^2_1}{\chi^2_2} \quad (31)$$

Como es el cociente entre dos variables aleatorias  $\chi^2$ , que son positivas, también será positiva, y su gráfica similar a la de  $\chi^2$



**Figura N° 11 - Distribución de Snedecor**

Su función de densidad surge del conocimiento de las variables aleatorias  $\chi^2$  que la integran, y es la siguiente:

$$f(F) = \frac{\Gamma\left(\frac{\nu_1 + \nu_2}{2}\right) \nu_1^{\frac{\nu_1}{2}} \nu_2^{\frac{\nu_2}{2}} F^{\frac{\nu_1}{2}-1}}{\Gamma\left(\frac{\nu_1}{2}\right) \Gamma\left(\frac{\nu_2}{2}\right) (\nu_2 + \nu_1 F)^{\frac{\nu_1 + \nu_2}{2}}}, \text{ para } F > 0 \quad (32)$$

Se utiliza fundamentalmente en la parte de inferencia estadística. Existen tablas que brindan valores de probabilidad del siguiente tipo:  $P(F > F_p, \nu_1, \nu_2) = P(\%)$ . Debido a que la función depende de los grados de libertad, se necesita una tabla a triple entrada para obtener valores tabulados de  $F$  que corresponden a distintas probabilidades y a distintos valores de los grados de libertad.

A veces es necesario tener en cuenta la siguiente relación:

$$F_{1-\alpha, \nu_2, \nu_1} = \frac{1}{F_{\alpha, \nu_1, \nu_2}} \quad (33)$$

### Distribución Log-Normal

Mientras que el modelo Normal surgía de la suma de pequeños efectos, también es importante considerar la distribución de un fenómeno que se origina como el resultado de un mecanismo multiplicativo, actuando sobre un número de factores.

Supóngase entonces que determinado proceso puede aproximarse por la siguiente expresión:

$$y_{t+1} = \alpha_t y_t + \beta_t z_t y_t = (\alpha_t + \beta_t z_t) y_t$$

Esto indicaría que valores sucesivos de una variable se pueden expresar a través del valor anterior más el valor de otra variable  $z_t$ , modificadas las dos por los parámetros  $\alpha_t$  y  $\beta_t$ , dependiendo del tipo de fenómeno y del instante de tiempo considerado.

Existe un gran número de sistemas físicos que pueden ser caracterizados por un mecanismo como el anterior.

$$y_t = \omega_t \omega_{t-1} \dots \omega_1 \omega_0 y_0 \quad (34)$$

Un ejemplo claro es el de los caudales medios diarios en una cuenca. El caudal de un determinado día se compondrá de una parte del valor anterior, debido a la persistencia provocada por el almacenamiento en el acuífero de la cuenca, mas otra parte que se origina por efectos externos, por ejemplo, una lluvia, que multiplicaría al valor anterior de manera tal que el valor resultante será mayor cuanto mayor sea el valor anterior y/o la humedad de la cuenca.

La variable de interés  $y_t$  se ha expresado en función de un gran número de variables, que pueden ser, cada una de ellas, difícil de estudiar y describir.

Aplicando logaritmos a ambos miembros de la expresión (34) se obtiene:

$$\ln y_t = \ln \omega_t + \ln \omega_{t-1} + \dots + \ln \omega_1 + \ln \omega_0 + \ln y_0 \quad (35)$$

Recordando nuevamente aquí el enunciado del teorema del límite central, puede decirse que el logaritmo natural de  $y$  se distribuye Normalmente, si cada uno de los logaritmos de las variables  $\omega_s$  cumplen con las condiciones establecidas para el modelo Normal.

La función de densidad del modelo log-Normal se obtendrá a partir de considerar que la variable aleatoria  $x$  observada cumple con la relación  $y = \ln x$ , siendo  $y$  distribuida Normalmente.

$$y \sim \text{Normal } f(y) = \frac{1}{\sigma_y \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{y - \mu_y}{\sigma_y} \right)^2}, \quad -\infty \leq y \leq +\infty$$

$$f(x) = f(y) \frac{dy}{dx}$$

$$\frac{dy}{dx} = \frac{1}{x}, \quad x > 0$$

$$f(x) = \frac{1}{x \sigma_y \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\ln x - \mu_y}{\sigma_y} \right)^2}, \quad x \geq 0$$

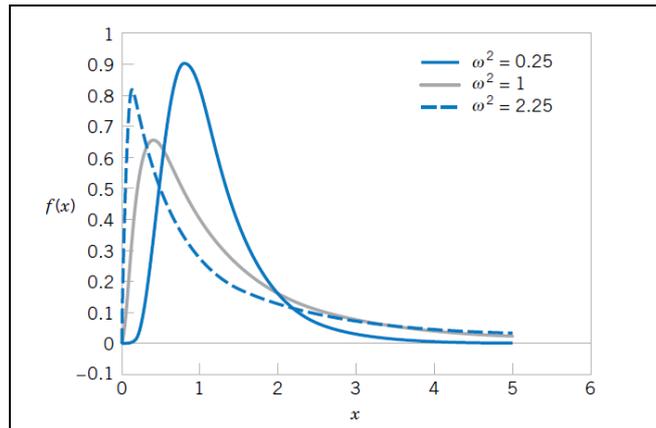
(36)

Sus parámetros son:

$$E(x) = e^{\left( \mu_y + \frac{\sigma_y^2}{2} \right)} \quad (37)$$

$$\text{Var}(x) = \mu_x^2 \left( e^{\sigma_y^2 - 1} \right) \quad (38)$$

En la siguiente figura se muestran distintas formas de este modelo para distintos valores de sus parámetros, apreciándose su gran flexibilidad. Por esto es que la transformación logarítmica de  $x$  tiene gran aceptación en diversas aplicaciones.



**Figura N° 12 - Distribución Log Normal**

Este modelo es adoptado particularmente para modelar caudales diarios, picos de descarga, crecidas anuales, precipitaciones diarias, mensuales y anuales.

Las tablas del modelo Normal pueden emplearse para evaluar probabilidades log-Normales, recordando que  $y = \ln x$ .

## MODELOS DE VALORES EXTREMOS

Los valores extremos ha constituido desde hace bastante tiempo una disciplina de gran interés, y no sólo para estadísticos sino, entre otros, para científicos e ingenieros.

Algunas aplicaciones de la teoría de valores extremos en la práctica, según distintos autores que además definen a la teoría de valores extremos como algo “curioso y fascinante” son: ráfagas de viento, contaminación en el aire, análisis de corrosión, inundaciones, sequías, efectos de aditivos en alimentos. Otros autores mencionan otras aplicaciones, como el estudio de la longevidad de la vida humana, la gestión de tráfico (en telecomunicaciones), la resistencia de materiales, la concentración de ozono, geología o meteorología (lluvias, vientos, etc).

La interpretación de lo que es algo “extremo” es complicada ya que su definición engloba varios atributos tales como “excepcional”, “sorprendente” y “catastrófico”. Según dichos autores, al ser como se ha dicho subjetivamente difícil definir a los valores extremos, es mejor caracterizarlos mediante, por ejemplo, sus propiedades estadísticas, observaciones, predictibilidad, mecanismos, etc.

La teoría de extremos para muestras está relacionada con el comportamiento límite del Máx  $\{x_1; x_2; \dots x_n\}$  o del Mín  $\{x_1; x_2; \dots x_n\}$  cuando  $n$  tiende a infinito.

En muchas situaciones prácticas en ingeniería, es de interés trabajar con el mayor o el menor valor de una variable aleatoria.

Si la variable  $Y$  es considerada el máximo de una serie de  $n$  variables aleatorias  $x_1 \dots x_n$  es posible obtener series  $y_t$  que estarán compuestas por los máximos anuales a

partir de valores diarios en  $n$  años. En cada año habrán 365 ó 366 valores de  $x$ , medios diarios, de los cuales se obtendrá el mayor para formar así la serie de  $n$  años de  $y_i$ .

Es posible obtener una expresión para la probabilidad de que el máximo sea menor o igual que un valor dado:

$$P(Y \leq y) = F(y) = P(\forall n \text{ de } x_i \leq y)$$

Si los valores de  $x$  son independientes, entonces:

$$\begin{aligned} F(y) &= P(X_1 \leq y)P(X_2 \leq y) \dots P(X_n \leq y) = \\ &= F_{x_1}(y)F_{x_2}(y) \dots F_{x_n}(y) \end{aligned}$$

Si los  $x_i$  son idénticamente distribuidos,  $F(x)$ , luego:

$$F(y) = [F_x(y)]^n$$

y, cuando  $n \rightarrow \infty$  : se deriva el modelo buscado.

De acuerdo a las características de la distribución inicial  $F(x)$  se originan tres tipos de modelos asintóticos de valores extremos.

Ya que  $F(x)$  no se conoce con exactitud, se estudiaron un conjunto de distribuciones asintóticas conocidas como Distribuciones de Valores Extremos, las que partiendo de ciertas funciones iniciales brindan un ajuste adecuado para grandes valores de  $n$ .

### **Modelo Tipo I:**

Surge de aquellas distribuciones iniciales que no tienen límite superior. El extremo de la curva correspondiente a la función de densidad debe decrecer tan rápidamente como una función exponencial; entonces, valores extremos provenientes de una distribución Normal, Log-Normal, Gamma, pueden ser ajustados por un Modelo Tipo I de máximos. En cambio, si la distribución inicial no está limitada en la dirección de los mínimos, se origina un Modelo Tipo I de mínimos.

### **Modelo Tipo II:**

Este modelo surge de aquellas distribuciones iniciales ilimitadas, y que poseen un número finito de momentos.

### **Modelo Tipo III:**

Este modelo surge cuando la distribución inicial está limitada en la dirección del valor extremo. Distribuciones iniciales del tipo Log-Normal, Gamma por ejemplo pueden ajustarse por un Modelo Tipo III.

Es muy usado en Hidrología, ya que es común que las variables estén limitadas inferiormente por el cero.

## Modelo Tipo I

### Distribución Gumbel

Existen una serie de condiciones que se deben establecer en el desarrollo del modelo:

- Las observaciones de las que se extraen los valores extremos deben ser independientes.
- Las observaciones deben ser hechas bajo las mismas condiciones, es decir, las distribuciones iniciales y los parámetros que contienen deben ser los mismos.
- El número de observaciones  $n$  de las cuales se extraen los valores extremos debe ser grande; en algunas situaciones, día y años son unidades naturales de periodicidad.

Se va a obtener la distribución del mayor de  $n$  valores de  $X_i$ , con  $n$  suficientemente grande. Supóngase que sólo se conoce que la distribución de los  $X_i$  no está limitada en la dirección positiva, y que su extremo superior decrece de una forma exponencial. Siendo esto así, puede expresarse la función de distribución de los  $X_i$  de la siguiente manera:

$$F(x) = 1 - e^{-g(x)} \quad (39)$$

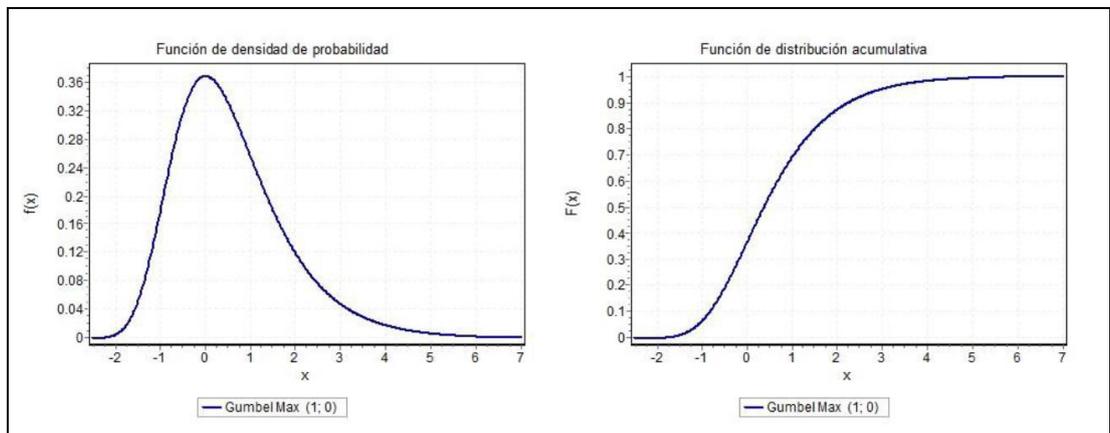
Luego, la distribución del mayor valor de los  $X_i$ , todos con distribución común correspondiente a la ecuación anterior, puede expresarse de la siguiente manera:

$$F(x) = e^{-e^{-\alpha(x-\mu_0)}}, \quad -\infty \leq x \leq +\infty \quad (40)$$

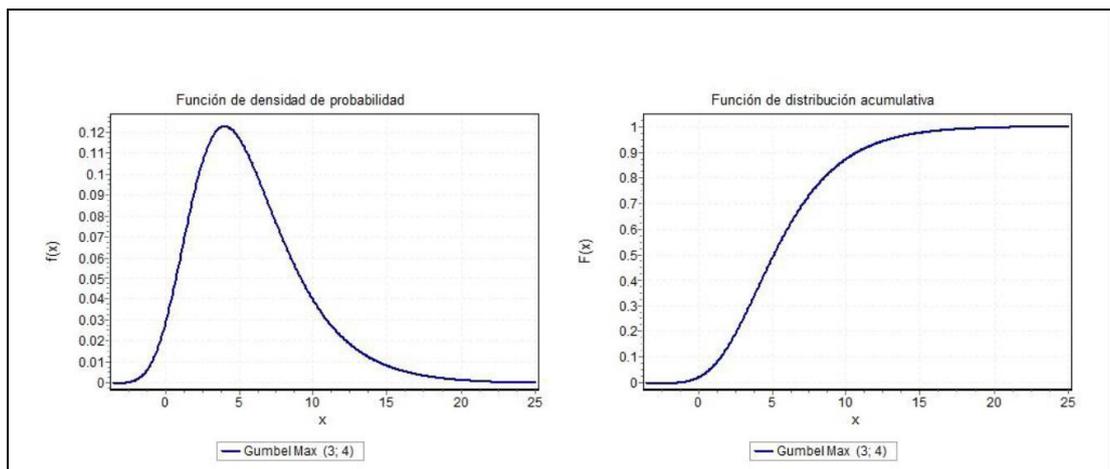
siendo, por lo tanto, su función de densidad la siguiente:

$$f(x) = \alpha e^{-\alpha(x-\mu_0) - e^{-\alpha(x-\mu_0)}} \quad (41)$$

siendo  $\alpha$  y  $\mu_0$  sus parámetros de escala y ubicación respectivamente.



**Figura N° 13** – Funciones de la distribución de Gumbel para distintos valores de sus parámetros de escala y ubicación



**Figura N° 14** – Funciones de la distribución de Gumbel para distintos valores de sus parámetros

Derivando e igualando a cero, se obtiene el máximo de la función, es decir el modo:

$$f'(x) = 0$$

$$\frac{d}{dx} \left( \alpha e^{-\alpha(x-\mu_0)} e^{-e^{-\alpha(x-\mu_0)}} \right) = 0$$

Recordar que  $\frac{d}{dx} e^x = e^x x'$

$$x = \mu_0$$

(42)

La media y la varianza:

$$\mu = \int_{-\infty}^{+\infty} x f(x) dx$$
$$\mu = E(x) = \mu_0 + \frac{0.577}{\alpha} \quad (43)$$

$$\text{Var}(x) = \sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx$$
$$\sigma^2 = \frac{\pi^2}{6\alpha^2} = \frac{1.645}{\alpha^2} \quad (44)$$

El coeficiente de asimetría es constante e independiente de los parámetros:

$$\gamma = 1.1396 \quad (45)$$

El que en las figuras 13 y 14 la variable sea asimétrica positiva no es casualidad, ya que al ser el coeficiente de asimetría para una variable con distribución de Gumbel para el máximo siempre positivo, cualquier variable con esa distribución será asimétrica positiva, independientemente de cuáles sean los valores de los parámetros.

Este modelo es especialmente utilizado para describir fenómenos tales como caudales máximos diarios en períodos de un año, picos de descargas horarias durante crecidas anuales, caudales máximos instantáneos anuales.

Este modelo puede ser aplicado también a valores mínimos:

$$F(x) = e^{-e^{\alpha(x-\mu_0)}} \quad (46)$$

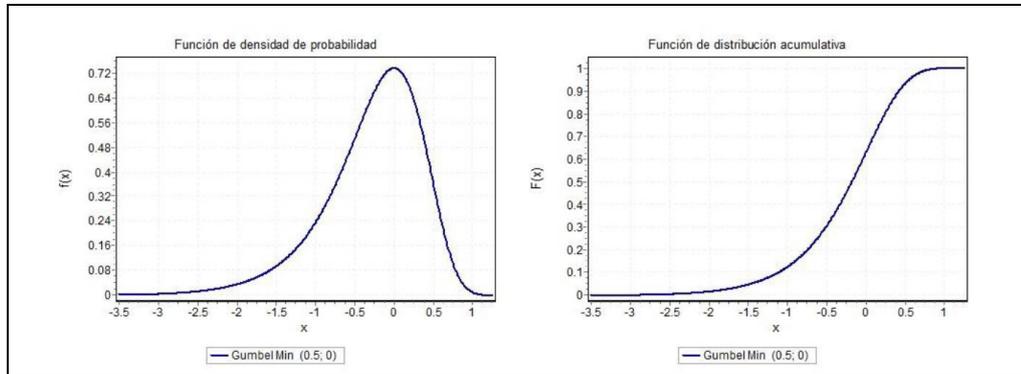
$$f(x) = \alpha e^{\alpha(x-\mu_0)} e^{-e^{\alpha(x-\mu_0)}} \quad (47)$$

con las características:

$$E(x) = \mu = \mu_0 - \frac{0.577}{\alpha}$$

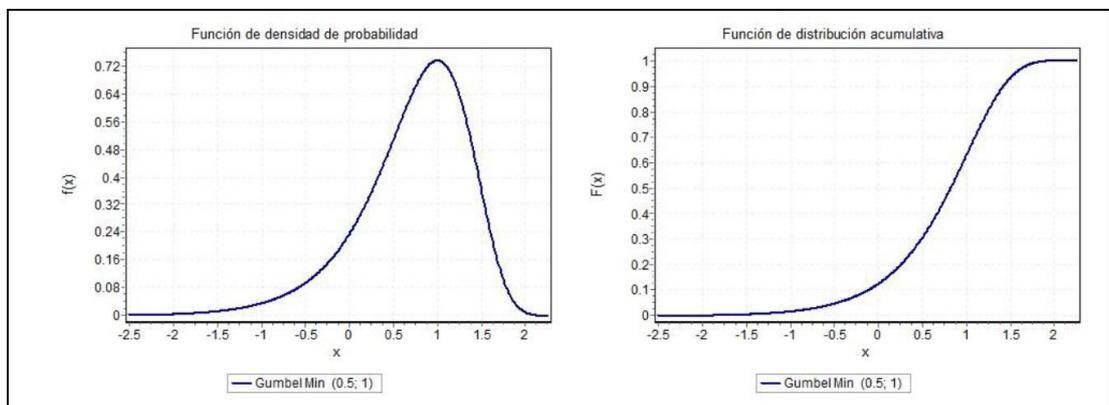
$$Var(x) = \sigma^2 = \frac{1.645}{\alpha^2} \quad (48)$$

$$\gamma = 1.1396$$



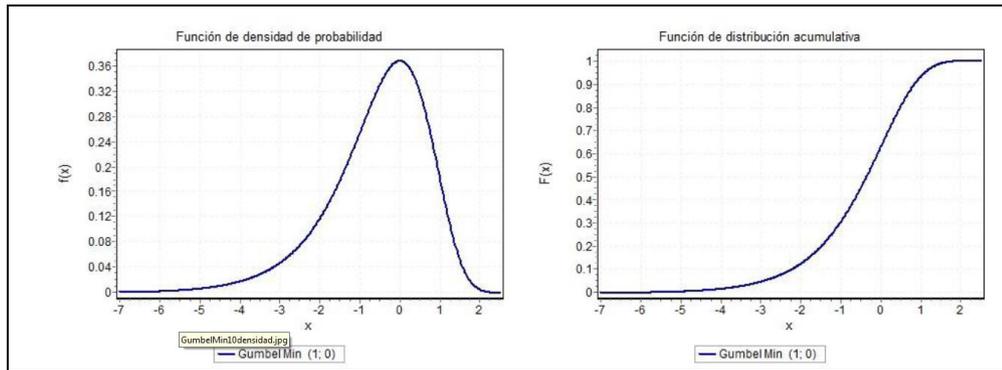
**Figura N° 15** – Funciones de la distribución de Gumbel para mínimos para distintos valores de sus parámetros

Se puede ver un cambio bastante evidente respecto a los gráficos vistos antes para la distribución de Gumbel para el máximo; ahora, las colas se concentran a la izquierda en la función de densidad, con lo que la distribución es asimétrica negativa. Igualmente, se puede ver que ahora la moda es 0, que coincide con el parámetro de localización como viene siendo habitual (el de escala vale 0.5).



**Figura N° 16** – Funciones de la distribución de Gumbel para mínimos para distintos valores de sus parámetros

Ahora se ha cambiado el parámetro de localización, la distribución se “traslada”; en este caso se mueve una unidad hacia la derecha, pues el parámetro de escala sigue siendo el mismo.



**Figura N° 17** – Funciones de la distribución de Gumbel para mínimos para distintos valores de sus parámetros

Se ha representado la distribución de Gumbel para el mínimo, con parámetro de localización igual a 0 y de escala igual a 1. Como en los dos casos anteriores, la cola está concentrada a la izquierda; la explicación de esto es que, opuesto al caso del máximo, para la distribución de Gumbel para el mínimo la densidad es asimétrica negativa, independientemente de los valores de los parámetros.

### Modelo Tipo III -Weibull

Este modelo se origina con las mismas consideraciones hechas para el anterior, a las cuales se les agrega las siguientes:

- La distribución de los  $X_i$  está limitada superiormente por un valor  $\omega$ .
- La distribución de  $X_i$  es del siguiente tipo:

$$P(X \leq x) = 1 - c(\omega - x)^k$$

con  $x \leq \omega, k > 0$

De esto se obtienen las funciones para el mayor valor de los  $X_i$ :

$$F(x) = e^{-\left(\frac{\omega-x}{\omega-\mu_0}\right)^k}, \quad x < \omega$$

$$f(x) = \frac{k}{\omega - \mu_0} \left(\frac{\omega-x}{\omega-\mu_0}\right)^{k-1} e^{-\left(\frac{\omega-x}{\omega-\mu_0}\right)^k} \quad (49)$$

Esta forma del modelo no es muy aplicada en muchas ramas de las ciencias, quizá por el límite superior que presenta. Una transformación de la variable resulta en el conocido modelo de Weibull. Si se trabaja con el valor negativo de la variable  $z = -x$ , estando  $x$  definida para valores menores que  $\varpi$ ,  $z$  estará definida para valores mayores que  $-\varpi$ . Llamando  $-\varpi = \varepsilon$ , y observando que las probabilidades de no excedencias serán ahora de excedencias, la función de probabilidad estará dada por:

$$P(Z \geq z) = e^{-\left(\frac{z-\varepsilon}{\mu_0-\varepsilon}\right)^k}$$

Por lo tanto, la función de distribución será:

$$P(Z \leq z) = 1 - e^{-\left(\frac{z-\varepsilon}{\mu_0-\varepsilon}\right)^k} \quad (50)$$

$$f(z) = \frac{k}{\mu_0-\varepsilon} \left(\frac{z-\varepsilon}{\mu_0-\varepsilon}\right)^{k-1} e^{-\left(\frac{z-\varepsilon}{\mu_0-\varepsilon}\right)^k}$$

Al trabajar con los valores negativos de una serie y maximizarlos, se estará obteniendo como extremo un valor mínimo, ya que el mayor valor de una serie negativa es el menor valor en valor absoluto.

Los parámetros de este modelo son  $\mu_0$ ,  $k$  y  $\varepsilon$ , siendo este último el límite inferior de los valores mínimos.

Las características son las siguientes:

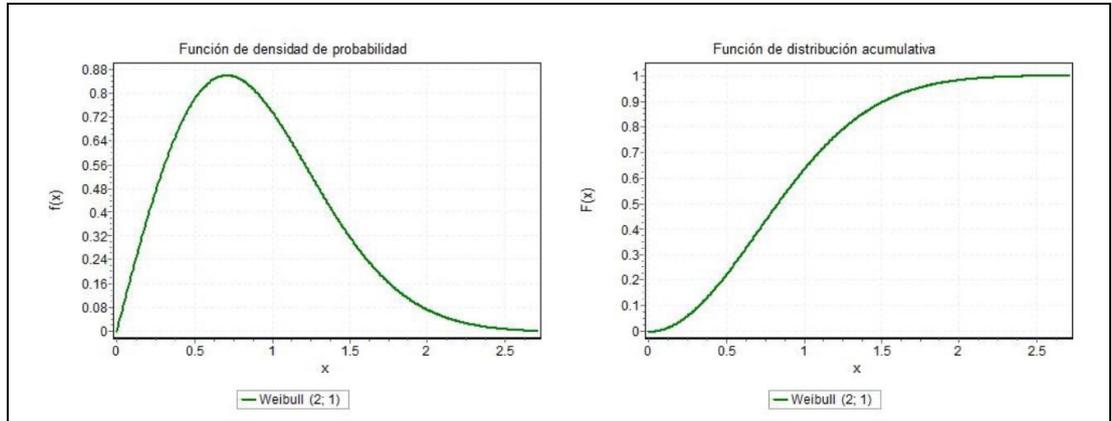
$$E(x) = \mu = \varepsilon + (\mu_0 - \varepsilon) \Gamma\left(1 + \frac{1}{k}\right) \quad (51)$$

$$\text{Var}(x) = \sigma^2 = (\mu_0 - \varepsilon)^2 \left[ \Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right) \right] \quad (52)$$

Generalmente es posible adoptar  $\varepsilon = 0$ , lo cual no es tan desacertado en el caso de variables hidroambientales, lo cual simplificaría las expresiones del modelo:

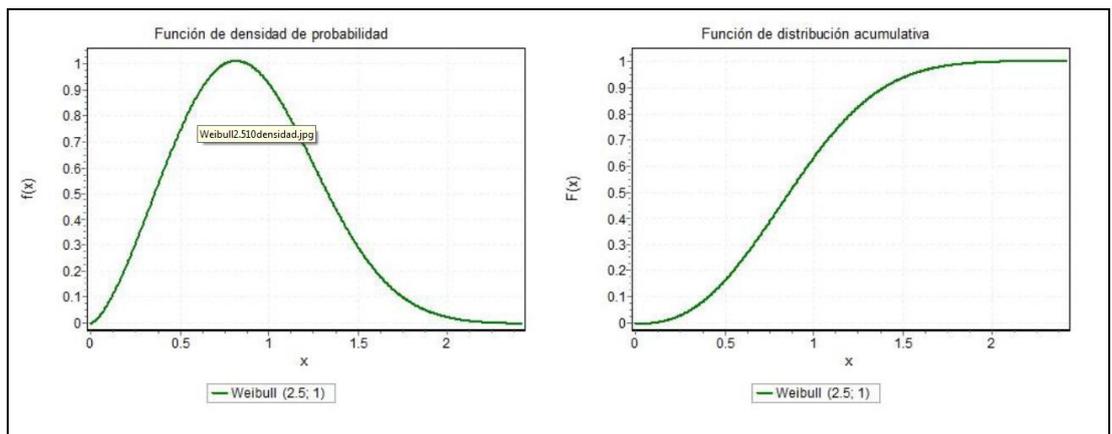
$$F(z) = 1 - e^{-\left(\frac{z}{\mu_0}\right)^k} \quad (53)$$

$$f(z) = \frac{k}{\mu_0} \left(\frac{z}{\mu_0}\right)^{k-1} e^{-\left(\frac{z}{\mu_0}\right)^k}$$



**Figura N° 18** – Funciones de la distribución de Weibull para distintos valores de sus parámetros

La primera distribución representada es la que tiene como parámetro de forma igual a 2 y de escala igual a 1; al no haber parámetro de localización, este se supone igual a 0, con lo cual la densidad y la distribución existen para valores mayores que ese valor. A la vista de la gráfica de la función de densidad, se puede deducir que en este caso la distribución es asimétrica positiva, mientras que de la representación de la función de distribución se puede deducir que el crecimiento suele ser constante hasta que  $x= 1.5$ , a partir de donde empieza a decaer ligeramente para crecer cada vez menos.



**Figura N° 19** – Funciones de la distribución de Weibull para distintos valores de sus parámetros

En este caso, el parámetro de forma vale 2.5, mientras que el de escala es igual a 1. La distribución es asimétrica positiva al estar la cola a la derecha y los valores con mayor probabilidad más a la izquierda.

Se observa que se tienen gráficas de forma flexible a medida que se cambian los parámetros con lo cual se adaptan a una buena cantidad de fenómenos tanto el modelo tipo I como el tipo III.